AIR FORCE INSTITUTE OF TECHNOLOGY

64-5

AIR UNIVERSITY
UNITED STATES AIR FORCE

SCHOOL OF ENGINEERING

WRIGHT–PATTERSON AIR FORCE BASE, OHIO

AF-WP-O-OCT 62 3,000

Best
Available
Copy

SEQUENTIAL RELIABILITY TESTS

APPLIED TO CHECKOUT EQUIPMENT

THESIS

GRE/ME/63-7 By Ben N. Theall

SEQUENTIAL RELIABILITY TESTS
APPLIED TO CHECKOUT EQUIPMENT

THESIS

Presented to the Faculty of the School of Engineering of
the Air Force Institute of Technology
Air University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

By

Ben Norman Theall, B.E.E.

Graduate Reliability Engineering

August 1963

Preface

This paper results from my efforts to understand the proper application of SPR (sequential probability ratio) tests, particularly as they apply to mean time between failure tests of checkout and similar equipment. My desire for such understanding grew from two beliefs. One, that a valid, physical test of an item is still one of the strongest tools we possess. Second, that I should be familar enough with a tool as strong as the SPR test to be able to tell a project engineer what the accumulated test hours buy.

Much of what I have learned on SPR tests is summarised in tables and appendices in the hope that they will be useful in practice. The Table I summary of changes expected in an SPR test as the test parameters change could prove convenient. The flow chart of Appendix J gives a suggested brute-force procedure for selecting a reasonable SPR test and attempts to relate various sections of this paper to the procedure.

I especially wish to thank Professor H. Kepler for his advice and encouragement during this recent time of trial. He diligently tried to prevent my striding off madly in all directions. I can only hope that I helped him succeed in that effort.

My wife also deserves a big share of any credit that this paper may create though none of the blame that my oversights may generate. Her tolerance of my recent disposition and work habits, and her help in preparing this paper have been signal contributions.

Ben N Theall

Ben N. Theall

Contents

Contents

Contents

List of Figures

List of Tables

List of Symbols

α	Producer's risk, P(Type I error), P(Reject H_0/H_0 true).
β	Consumer's risk, P(Type II error), P(Accept H_0/H_1 true).
A	One of the decision boundaries in SPR test. See Eq 2.
B	The other decision boundary in SPR test. See Eq 3.
AFN	Expected number of failures ($E_\theta(r)$ in Handbook H108 (Ref 18)).
E(t)	Expected calendar waiting time for a decision.
H_0	The null hypothesis that $T = T_0$ in this paper. (In effect $T > T_0$)
H_1	The alternate hypothesis that $T = T_1$ in this paper. ($T < T_1$)
k	Discrimination ratio T_0/T_1 (note, reversed from test ratio here)
L(T)	Probability that the lot will be accepted given T true.
MTBF	Mean time between failure, T in this paper.
n	Number of items on test or the sample size.
O.C.	Operating characteristic. L(T) versus T is the O.C. curve.
P(x/y)	Probability that x will occur given y is true.
r	Number of failures.
R	Reliability or probability of no failure in the time considered.
SPR	Sequential probability ratio.
t	Time of operation, mission, test, etc.
T	Mean or average time between failures.
T_0	Minimum T of lot if H_0 is true. (Termed θ_0 in Handbook H108, termed T_1 in AGREE (Ref 13)).
T_1	Maximum T of lot if H_1 is true. (Termed θ_1 in Handbook H108, termed T_2 in AGREE).

List of Symbols

T_1/T_0	Test ratio, (termed Θ_1/Θ_0 in Handbook H108).
$V(t)$	Sum of test time on all equipments.
$V(T)$	Expected sum of test time on all items if T is true.
$V(T)/T_0$	Expected sum of test time on all items if T is true. Expressed in multiples of T_0.

Abstract

This paper briefly discusses the need for reliability test of mission support aerospace ground equipment (AGE) such as checkout equipment. The discussions are aimed primarily at the sequential probability ratio (SPR) test for mean time between failure (termed T in this paper) determination when the distribution of failures can be assumed exponential. Other distributions are only mentioned. No specific checkout equipment is considered.

Summary tables and charts are included in several areas in the hope that they will prove useful in practice for selecting an acceptable SPR test. The following subjects are discussed: considerations in setting acceptable test risks, conditions under which SPR tests are appropriate, the theory of SPR tests and the resulting relationships used in SPR tests of T, and the expected qualitative variations in an SPR test when different test parameters are varied. It is pointed out that the expected failure number (AFN) characteristic yields incentive for the contractor to submit equipment with T well above the two values of T, T_1 and T_0, selected as the decision limit parameters of the test. As the true T increases above those test-establishing parameters, the AFN and the expected waiting time decrease.

It is concluded that Handbook H108, Sampling Procedures and Tables for Life and Reliability Testing (Based on the Exponential Distribution), reflects much of Epstein's work on SPR tests and, as such,

is an excellent reference for practical use. Data from that handbook are manipulated to indicate that for specific assumptions, a simple exponential relationship may be established between the expected test time and the ratio of the two test establishing parameters T_1 and T_0. It is indicated that such a relationship could be used in conjunction with cost considerations to lead to a simple method for selecting the T_1/T_0 ratio yielding lowest cost. The development of such a method is not completed in this paper. A flow chart suggesting one iterative procedure for selecting an appropriate SPR test for a particular problem closes the paper.

SEQUENTIAL RELIABILITY TESTS APPLIED TO CHECKOUT EQUIPMENT

I. Introduction

The importance of the probability that an item will work when it is needed depends on what the item is and where it is being used. In military equipment used in combat that probability or reliability is very important. Usually, the probability of success should be high to save both lives and dollars, but whether the probability is high or, due to technical or financial constraints, is low, the activity responsible for operational planning should know what some measure of that reliability is. Proper operational use of even low reliability items can lead to acceptable probability of accomplishing a particular mission. Knowledge of the true item reliability helps in developing the proper operational use procedures.

Aerospace Ground Equipment (AGE) such as equipment used to check-out an aircraft or missile before a mission can certainly exert a strong influence on the probability of mission success. Therefore, some measure of the AGE reliability should be known to allow for proper operational use and planning. The most commonly used measure of reliability is the item's MTBF (mean-time-between-failure), termed T in this paper. One way to measure the T or to assure that it is above some specified value is to use an SPR (sequential probability ratio) test to test the hypothesis that the T of an item is at least

the specified number. As discussed in the sections following, the SPR test is usually the fastest and the least expensive test for a particular risk level. This then leads to the purpose of this paper.

Purpose

The purpose of this paper is as follows:

1. Locate or develop the background information needed to write step-by-step procedures to be followed as a guide in selecting a reasonable SPR test of T (MTBF) of checkout or other equipment.

2. Find and list what should be considered in selecting the risks involved in an SPR test.

3. Obtain information on SPR tests sufficient to allow explaining to equipment project engineers what the tests can do for their equipment.

4. Determine the effects of varying the different parameters of an SPR test.

Scope

This paper shall cover the following general areas:

1. SPR tests for T when the equipment failure distribution is exponential or governed by a constant hazard rate. Other tests and distributions are mentioned but are not covered.

2. The general situation assumed is that in which a reliability engineer or project engineer is requested to formulate the reliability test provisions for a specific type of checkout equipment or other equipment being procured by separate contract, not

as part of a system contract. An alternative to this is that the engineer is requested to analyze and comment on a contractor-proposed test plan for reliability demonstration.

3. To avoid misunderstanding, it should be made clear that no details of any particular checkout equipment have been included. A piece of checkout equipment used in a critical application was kept in mind to direct the information and procedures developed. Some characteristics of checkout equipment that should be considered in reliability tests are considered in Appendix H.

Limitations

Longevity or wear-out tests are not considered. The test period is assumed to remain in the constant hazard rate or flat portion of the traditional bathtub curve.

The very important area of environmental testing is not covered with the exception of the list of reminders in Appendix H. The accelerating effect of stringent environments on reliability tests should be a boon to reliability activities when data giving correlation factors becomes available, but no information on such factors was found for this paper.

Assumptions

The following subparagraphs 1 through 10 list assumptions that apply in this paper and to the procedures described in this paper:

1. The hazard rate is constant, Carhart was one of many indicating that the assumption is valid for much electronic and

electromechanical equipment, but that the assumption bears questioning each time it is made (Ref 2:120-125).

2. Testing is with replacement or with early repair of failed items and return to test. This is the normal procedure in reliability testing by SPR tests since the waiting time is expected to be shorter than in the nonreplacement case.

3. Reducing test and waiting time is a major consideration. Lloyd and Lipow indicate that a nonreplacement test can not cost more than a replacement test although the replacement test should require less test time (Ref 12:313). It is assumed that any cost difference will be more than offset by the earlier decision and earlier equipment delivery resulting from a test.

4. The reliability and the mission time have been directed or established or can be estimated by the activity responsible for the equipment.

5. The items making up the sample will be representative of the entire lot submitted for acceptance. If there are appreciable time lags in production between items or if there are design or process changes from one equipment to another, this assumption should be kept in mind.

6. Contract quantities will be so small that production sampling techniques will be difficult or impossible to apply. The number of items available for test are limited, and an accept or reject decision must be made based on the data and the time available.

7. Usually the number of items that can be made available for reliability testing is limited - every one available will be used. If n, the sample size, can be varied, then that variability should be added to the flow chart of Appendix J in the search for a satisfactory test plan.

8. Adequate test procedures and facilities have been developed, are available, and will be used during reliability testing. A test to demonstrate reliability must have a rational relationship to the expected use of the item. This is a very important assumption and deserves attention in writing the reliability test provisions. A generator may have a very long time-between-failure characteristic when it is tested as a doorstop.

9. One of the major tasks of reliability personnel is to sell project engineers on the value of reliability tests. Directives and laws, as in prohibition, civil liberties, and the like, may force the motions of reliability principles and practices being applied. But that application will not be fully effective until the person who should be best able to pinpoint the potential trouble areas, the project engineer, is sold on the value of reliability tests to his equipment.

10. It is further assumed that the reader of this paper has at least a moderate knowledge of reliability practices, is interested in using SPR test to save test time, and desires to have information to help convince project engineers that SPR tests are more than meticulous mensurations by mendacious magicians.

Need for Test

Complete reliability programs require test or demonstration of a quantitative measure of reliability and a formal reliability organisation with adequate procedures and practices all monitored by the buyer. This, naturally, brings up the frustrating question of how small should the organisation be for a large contractor and how large should it be for a small contractor? The answer to that is well beyond the scope of this paper.

Regardless of the size of the reliability organisation in the contractor's plant, one rule is valid; a requirement without a valid test or demonstration is not a requirement - it is a request. The requirements section of a specification or exhibit may call for grand reliability, but it will mean little if the test section does not include a good test for the requirement.

The contractor will sensibly design and produce primarily to pass the tests that control whether his product is accepted or rejected. No matter what the process and practice controls are, the crucial point is the acceptance test. Process and practice controls should be used as substitutes for a physical test of the item only when a suitable test or test facility is not available and when it is truly known what processes and practices insure the item characteristic desired. Certainly, few people claim to know exactly what processes and practices insure reliable equipment in most areas. Experience in bridge building and structural engineering has indicated the safety

factors that reduce the probability of failure of many stationary structures to low levels, but such knowledge is not yet available for many Air Force procurements where reliability is still a major concern. Therefore, whenever design and process controls are not adequate to insure the reliability needed, a reliability test must be conducted.

Reducing Test Costs. If two different test plans give the same protection against errors of judgement or against wrong decisions it is an engineer's responsibility to choose the less-expensive test if all other considerations are equal. This leads to SPR tests. As described later, SPR tests usually save time over conventional tests.

SPR Test History

Wald devised sequential analysis in March and April 1943 to meet the need for methods of obtaining reliable conclusions from limited data. The method was for application to problems of analyzing combat experience, military operations, and development problems. At that time the method was classified due to its high value. The classification was dropped in 1945 (Ref 14:1). Since that time much effort and literature has been devoted to sequential tests, or SPR tests as they are commonly referred to. Notable in the efforts on SPR tests, particularly for life testing under the exponential or Poisson case, has been the work of Professor B. Epstein. Sequential tests are a natural development starting with single sampling plans, to double sampling plans such as those of Dodge and Romig, to multiple sampling plans,

and then to SPR plans allowing for something approaching a continuous probability of a decision to accept or reject.

Method

The following general approach is used in this paper: what considerations influence the risks, when are SPR tests appropriate, what is the basic theory of SPR tests of exponential distributions, what does analysis show the test parameter interactions to be, and what is a reasonable procedure to follow in selecting an SPR test?

II. Sampling and Risks

The problems of primary interest are those involved in drawing inferences about the true character of a population when only limited data are available. One particular sort of distribution of failures is assumed, but the problems and risks involved are similar in all attempts for such inferences. In reliability testing it is usually, as here, assumed that the distribution of failures obeys an exponential law or, more specifically, has a constant hazard rate. If true, the distribution is governed by only one parameter, T.

In setting the sampling risks and the magnitudes of the related parameters to be estimated, the basic considerations are similar for any test plan for testing hypotheses. The magnitudes must, in general, be set by the engineering or the operational problem involved. The risks must be set by study or judgement of the penalties to be incurred if errors of decision are made. Setting the values of the hypothesis to be tested and selecting the acceptable risk levels concern reliability and project engineers, whether in preparing a test plan to be included in a contract or in analyzing a contractor-proposed test plan. Some of the basic considerations appropriate in selecting such test-plan characteristics are discussed in the next few sections.

Preferences

As Wald points out the perfect test would yield an operating characteristic similar to that shown in Figure 1 (Ref 15:27). If T_c were some value of a characteristic T of the items to be tested, the test would reject all lots with T less than T_c and would accept all lots with T greater than T_c. As mentioned previously, T is MTBF here, but, in general, T could be any characteristic such as tensile strength. Even if the buyer could establish a number for T_c very accurately, the test for such high discrimination would require 100 % inspection and even that much effort would probably not yield an ability to discriminate between good and bad lots as sharp as that shown in the figure. Human errors alone would force the vertical line to some finite slope (Ref 17:3).

In specifying a particular characteristic there will usually be zones of preference and degrees of preference to accept or reject items of various T values. In Figure 1 the buyer would prefer to accept all lots to the right of T_c and to reject all lots to the left. Figure 2 indicates that the preference to reject or accept would increase as the actual T (T true) deviation from T_c increases. Perhaps T slightly below planned (T_c) would not reduce system effectiveness much. The preference to reject would be low. If T were much lower than T_c the lot might be worthless or detrimental to the buyer and the preference for rejection would be very strong. The same sort of reasoning would apply to the high T side with strong preference for acceptance

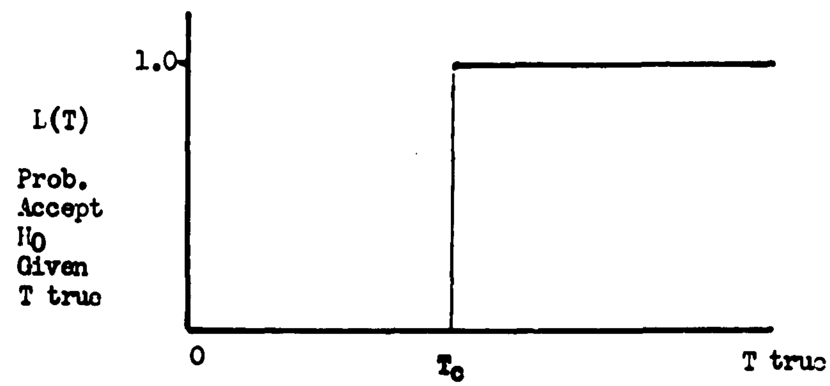


Fig. 1

Perfect O.C. Curve

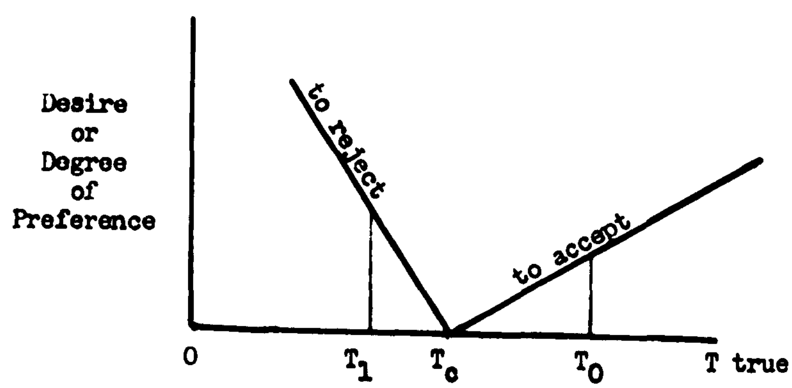


Fig.2

Reject and Accept Preferences

on items appreciably better than T_c. Probably neither preference line would be linear and there would be no necessity that they be the same shape, but the trends should be as indicated.

Setting Risks. A sketch similar to Figure 2 might help the buyer in establishing the α and β risks. For some particular T_1 below T_c a preference to reject could be set. This could be the result of some sort of trade-off or economic study, but as H105, Administration of Sampling Procedures for Acceptance Inspection, points out, dollar values are often difficult to set for such preference points (Ref 17:9). The preference to reject at a specific T_1 will primarily be based on the judgment of the buyer concerning the importance of the item and the amount of trouble to be expected if lots accepted are as bad as T_1. At some point T_0 a preference to accept can be estimated, primarily on judgement again.

The two preference points can be converted to a form of an O.C. curve as sketched in Figure 3. At point T_1 the preference to reject can be called $1-\beta$ where β is the probability that lots with T actually equal to T_1 will be accepted. $1-\beta$ will be the probability that such a lot will be rejected. This could be considered a form of confidence statement in that the buyer would be sure that no more than 100 (β) % of all the lots accepted would have a T actually less than T_1. At T_0 the preference to accept can be called $1-\alpha$, the probability that a lot with T equal to T_0 will be accepted. The contractor will have quite some interest in this point since α is the probability that

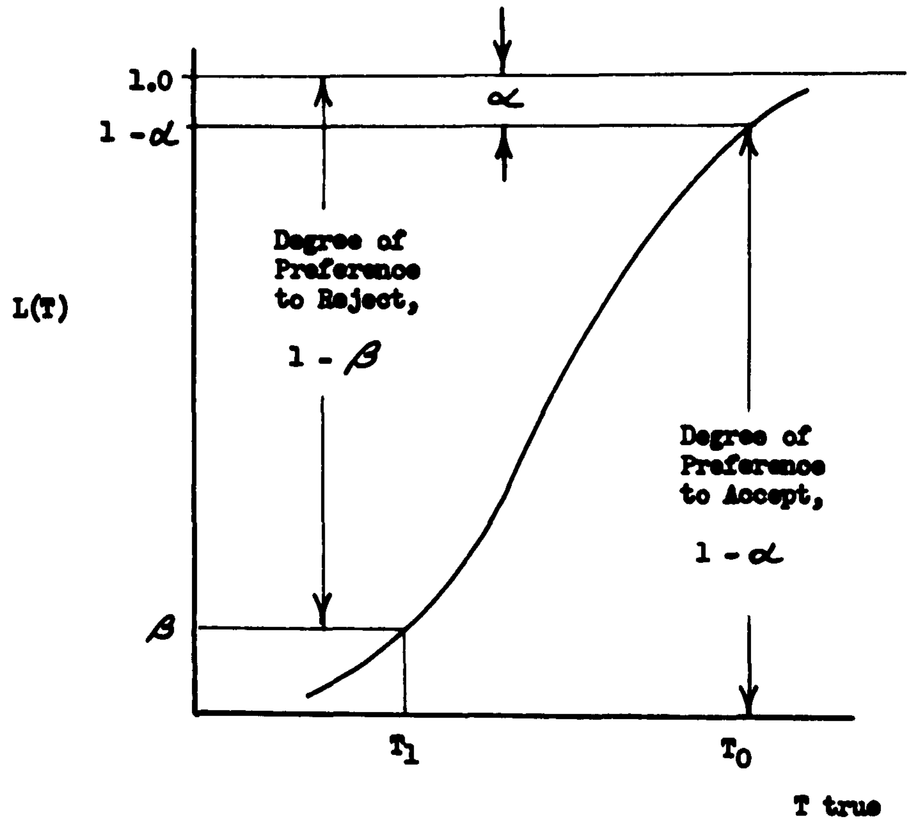


Fig. 3

Risks α and β in O.C. Curve

that the same lot will be rejected. If the test were to assure accepting lots with low characteristic, for example if the abscissa represented fraction defective or probability of failing before some specific time, the slope of the O.C. curve would be reversed.

T_1 and β Relationship

Going to a higher T_1/T_0 ratio or to a lower β risk improves the discrimination of the test. That is it improves the test's ability to separate good from bad lots. It seems that the same β risk should not apply to all sorts of equipment - some are more critical than others. However, if it is desired to retain the normally accepted 0.1 for β for equipments of different criticality, the O.C. curves can be used to represent many T_1, β points as some equivalent T_1, β equal 0.1 point. T_1, the low T value that is to be accepted only β portion of the time, can be changed in combination with β as discussed below.

Referring to Figure 4, say some particular combination of T_1 and β is being considered, and it has been found that the expected test time or costs are too high. The curve O.C.1. describes the acceptance probabilities associated with the test. Decreasing the slope of the O.C. curve from O.C.1 to O.C.n will decrease the expected test time. As shown, if a risk of accepting T_1 items as high as β_n can be tolerated, T_1 and the new β_n describe the two lower points for the new test described by O.C.n. However, that same O.C.n test can just as well be described by the original β and a new lower $T_{1,n}$ read from

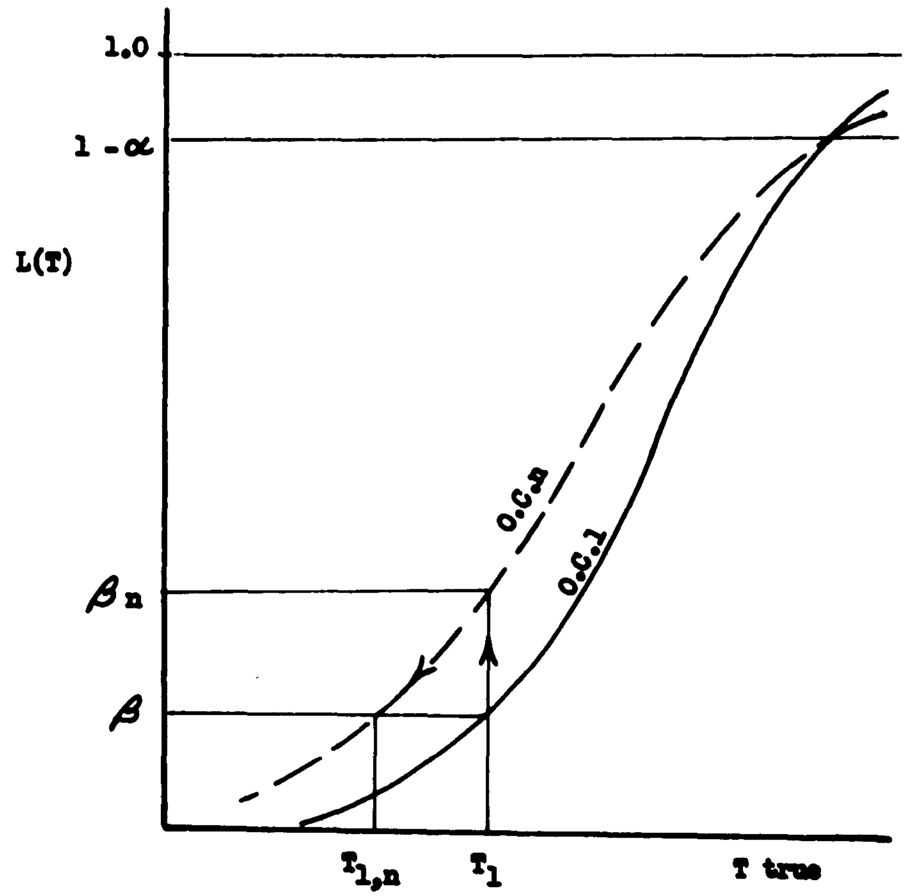


Fig. 4

Interchange of T_1, β Pairs

from the O.C.n curve. Useful O.C. curves can be found in Handbook H108, <u>Sampling Procedures and Tables for Life and Reliability Testing (Based on the Exponential Distribution)</u>, (Ref 18:2.13-2.16).

<u>β Held at 0.1</u>. If an estimate can be made of the lower T_1 which the buyer can tolerate, but the associated β is other than the commonly used 0.1, an equivalent T_1 at β equal to 0.1 can be found from the O.C. curves as described above. The new T_1, β can yield a test equivalent to that desired. Also, if it is desired to speak only of a high numbered T_1 for a particular test, β can be set high.

<u>Sampling Risks</u>

The spread of T_1 from [illegible] and the associated α and β acceptable risks allow sampling plans to be used. If the population T is actually T_1, only β of all such lots will be accepted on the average. If the T is actually T_0 , only α of such lots will be rejected on the average. When T_1 and T_0 are very close and the related preferences for rejection and acceptance are very small the test must approach the perfect O.C. curve. Such a test would be very costly unless the items and the tests were very simple. If testing is destructive such a sharply discriminating test would allow but few items delivered - most of them would be destroyed in testing. Therefore, a test plan must be selected that is, as most things on this earth, some compromise between the desire for low risk of error and the costs of getting low risk. Appendix F lists some of the points to be considered in choosing or accepting the α and β risks for a particular test.

Selecting the Sort of Test

The considerations of risk discussed above apply to any sort of test involving sampling. They apply to single sampling tests such as those described by MIL-STD-105, Sampling Procedures and Tables for Inspection by Attributes, and to the life tests terminated at a preassigned time or at a preassigned number of failures such as those described in the first sections of H108 (Ref 18). Tests such as these are termed conventional tests in this paper. The considerations also apply to SPR tests with the difference that often the SPR test will give the same risks for less inspection or time. Figure 5 illustrates the expected reduction in calendar waiting time for a decision when an SPR test is used rather than a conventional preassigned time test. In general, the SPR test is preferable.

In testing only for catastrophic failures that are not self-repairing a preassigned time test could be preferable to an SPR test. In a preassigned time test the number of measurements made during the test time could be reduced to a few or even to only one measurement at the end of the preassigned time. If each measurement to be made is complex and expensive, the measurement costs of the preassigned time test might be reduced sufficiently to offset the shorter expected waiting time of an SPR test. Measurements during the test are mandatory in SPR tests.

However, in items such as checkout equipment an error that is to be counted as a failure may occur often during the test period but may

Note: O.C. curve for each plan gives
T_0 of 1500, α of 0.05
T_1 of 300, β of 0.10
Approximately Code B-4 of Handbook H108,
T_1/T_0 of 0.205.
Both tests below are with replacement.

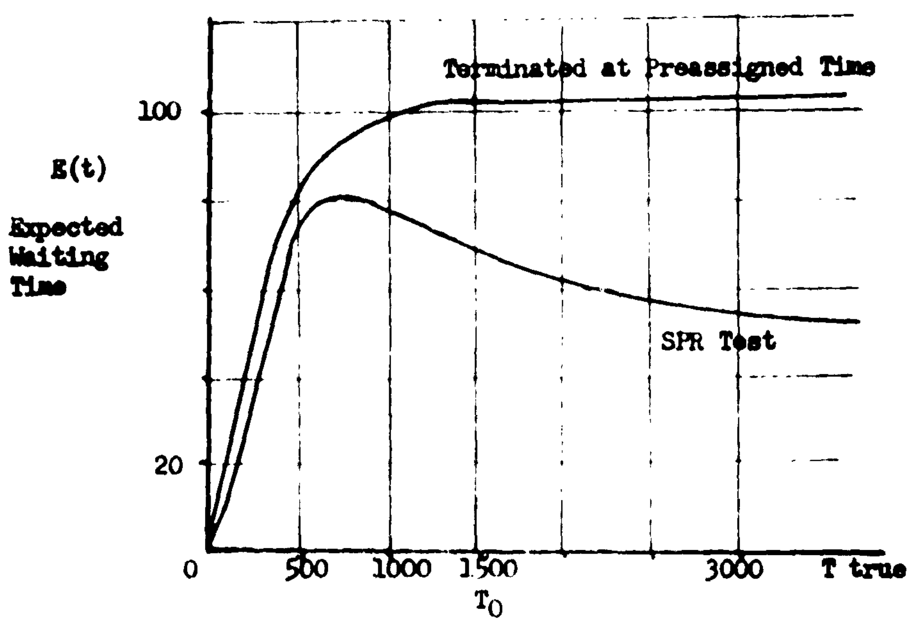


Fig. 5

SPR Test Savings in Test Time

(From Ref 18:2.68)

not occur at some preassigned test time. Therefore, a preassigned time test may miss many errors that occur during the test period. Measurements during the test period must be made at some frequency related to expected service use or must be made continuously if practicable. For such measurements an SPR test will give a shorter expected waiting time and measurement cost than an equivalent strength preassigned time test.

If the measurements to be made are simple and inexpensive, any advantage a preassigned time test might have becomes less. Appendix G lists other considerations in selecting a test.

III. SPR Tests

The sequential probability ratio test is usually an excellent method of obtaining a specified amount of information with a minimum amount of testing. The SPR used is particularly appropriate when the exponential distribution can reasonably be assumed since the one population characteristic, T, simplifies the test.

Developing the SPR

Appendix A gives the procedure for developing the SPR and the decision criterion to be used. The material there is similar to that found in the literature with some clarifications. Appendix B illustrates the development of the A and B (decision boundaries of the basic SPR inequality) and α and β relationships in the decision criteria. Specific SPR tests can be developed for an infinite number of combinations of α and β risks, and T_1/T_0 test ratios if desired. Those three parameters, the two risks and the test ratio, are sufficient to completely describe a test. The need for great detail in specifying the parameters to hairline differences from existing test plans is questionable unless large lots of continuous production are involved. H108 (Ref 18) gives many test plans for a wide selection of α and test ratios. In those plans β is usually set at 0.1, but as explained earlier, the O.C. curves can be used in conjunction with changes in T_1 to give a wide range of β risks. It is strongly recommended that H108

plans be used whenever possible to simplify the calculations needed and since H108 gives the related O.C. curves.

Time as the Variable

SPR tests are not limited to mean time tests but can be applied to many tests of hypotheses. The distributions need not be exponential. Wald and the Columbia Statistical Research Group, with which he was associated at the time, give the procedures for developing SPR tests for several different sorts of hypotheses and distributions (Ref 14 and 15). For example, an SPR test could be used in a test of the mean pull-in voltage of relays. One relay at a time would be selected and increasing voltage applied until the contacts closed. The variable would be pull-in voltage and would lead to a decision about mean voltage per unit. The unit of test would be one relay. As the inspector began applying voltage to a relay in the sequence there would be some expected (mean or average) voltage at which that unit would, on the average, pull in.

In most reliability SPR tests the variable in question is not voltage or length but is the time to fail. The variable measured is the time to fail which leads to a decision about the mean time to fail. The unit of test can be considered one failure, not one equipment or relay. After any particular failure there is some expected test time to be accumulated before, on the average, the next failure will occur. Therefore, SPR tests can be used to demonstrate mean time to fail as well as average break-down voltage, average cures per treatment, and other population characteristics.

Normalised Decision Relationship

Eq 1 below is the basis of this discussion and of most of the work on SPR tests of exponential distributions. The relationship is normalised to T_0, the high target T. It could just as well have been normalised to T_1, but since the consumer can certainly hope that the test will be operating in the T_0 area, the normalisation used, which is that commonly used, seems reasonable. The background theory and the absolute relationships leading to Eq 1 can be found in Appendix A.

$$\frac{-\ln A}{(T_0/T_1)-1} + \frac{r \ln (T_0/T_1)}{(T_0/T_1)-1} < \frac{V(t)}{T_0} < \frac{-\ln B}{(T_0/T_1)-1} + \frac{r \ln (T_0/T_1)}{(T_0/T_1)-1} \tag{1}$$

(Accept H_1 side)
(Reject H_0)

(Accept H_0 side)
(Reject H_1)

where

$$A = \frac{1-\beta}{\alpha} \tag{2}$$

$$B = \frac{\beta}{1-\alpha} \tag{3}$$

α = Producer's risk = P(Reject H_0/H_0 true)

β = Consumer's risk = P(Accept H_0/H_1 true)

T_0 = T at which α of such lots will be rejected $\doteq$ high target T.

T_1 = T at which β of such lots will be accepted = low target T.

r = Sum of failures observed to the time being considered.

V(t) = Sum of test time on all items under observation.

and

T_1/T_0 = Test ratio. (Note inversion in Eq 1.)

and in effect

$$H_0: \; T \geq T_0$$
$$H_1: \; T \leq T_1 \qquad (T_1 < T_0) \tag{4}$$

Eq 1 is commonly written

$$\frac{h_1}{T_0} + \frac{r\,s}{T_0} < \frac{V(t)}{T_0} < \frac{h_0}{T_0} + \frac{r\,s}{T_0} \tag{5}$$

Decision Criterion. If the inequality of Eq 1 is violated on the left side the hypothesis that the lot T is equal to or greater than T_0 is rejected (H_0 rejected) and H_1 is accepted. If the right side inequality is violated, H_0 is accepted. If neither side is violated, that is if the normalized sum of all the equipment test time up to the time being considered stays within the inequality signs, no decision can be made at the risk levels used. The test must be continued. Before the next failure, V(t) could increase enough to violate the right side inequality and the lot could be accepted, but a failure would be required before a reject decision could be made. This can be seen in the figure that follows. As time accumulates on the test the total test time in multiples of T_0 is recorded. Each chargeable failure is recorded and a decision attempt made at each failure and as often between as desired.

<u>Decision Methods</u>. The accumulated time and the accumulated failures can be compared to tables developed from Eq 1 or they can be plotted graphically. Eqs 1 and 5 describe two linear and parallel lines of the form $y = ax + b$. Figure 6 indicates the sort of decision lines to be expected based on those equations. AMCP 74-1 gives a similar plot for the AGREE Task Group 3 test (Ref 16:8). Each line will have slope s/T_0 when $V(t)/T_0$ is plotted against r. The accept or high intercept on the ordinate (time axis) will be h_0/T_0 and the reject intercept h_1/T_0. Example plots of two test histories are shown. After each failure, increasing test time with no further failure is indicated by a vertical line corresponding to the added time. When a failure occurs, the event plot jumps to the right in discrete steps describing the accumulated r. The decision lines are not plotted in discrete steps. A decision can be made at any point on either one of the decision lines where a test history or event plot happens to touch. AGREE (Ref 13:89,136-137) and MIL-R-26667 (Ref 19:38-39) use normalized tables. Either method of tables or plots is satisfactory since both represent the same decision criteria.

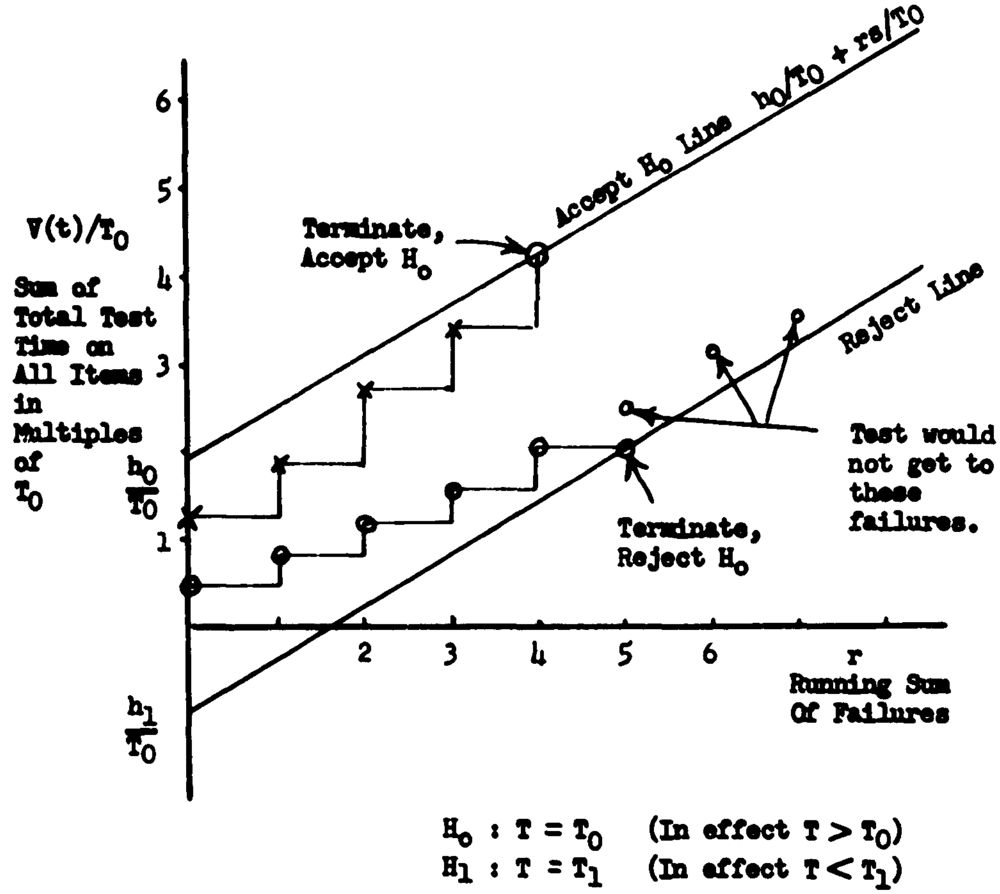


Fig. 6

Typical SPR Test Decision Plot

IV. Interaction of Test Parameter Variations

What is the influence on the test described by Eqs 1 and 5 and Figure 6 of changes in the α and β risks and in the T_1 and T_0 parameters? There obviously is not one grand formula that will solve every problem that may be encountered. Probably derivatives of Eq 1 could be taken with respect to the various parameters to determine some sort of sensitivity of the test to various changes. This would become very difficult to handle when the O.C. curve calculations are made. The parametric equations causing this difficulty will be mentioned later. The next sections analyze the influence of parameter changes. The distribution is still assumed exponential except where specifically considered as another form. For the reader not particularly interested in the detailed analysis of this chapter, Chapter VI includes a summary of the results of this chapter in what is hoped is a convenient table.

Variation of Slope s/T_0 with T_1 and T_0

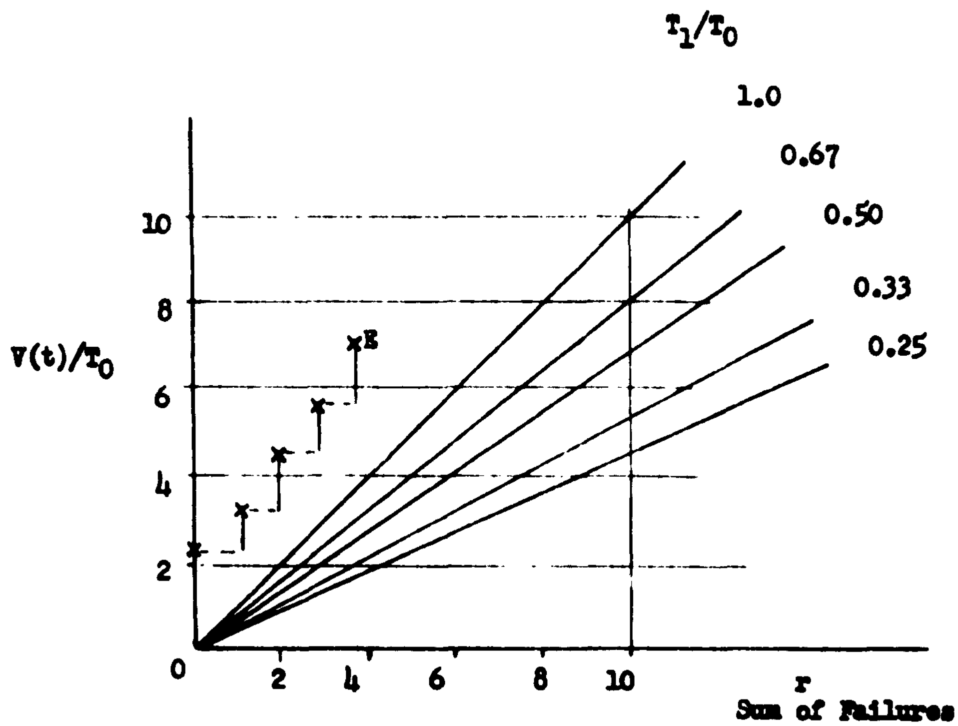
The slope of both decision lines is s/T_0. From Eqs 1 and 5

$$\frac{s}{T_0} = \frac{\ln (T_0/T_1)}{\frac{T_0}{T_1} - 1} \qquad (6)$$

Therefore, the slope of the lines is independent of the α, β risks. Calculating s/T_0 for a few representative test ratio, T_1/T_0, values leads to Figure 7. As the test ratio increases the decision lines are rotated up toward the accept region. It should be noted that the values obtained agree with Handbook H108 tables (Ref 18:2.64). As the test ratio decreases the slope becomes less, and as T_1 approaches zero the slope approaches zero meaning the probability of rejecting any lot decreases.

As the test ratio increases, the probability of accepting H_0 decreases for any particular sequence of events (E in the figure for example) and the test is more stringent. Also as the test ratio increases, T_1 approaches T_0. Taking the derivative of both top and bottom of Eq 6 with respect to T_0/T_1 to obtain an estimate of the limit, it is noted that the limit becomes T_1/T_0 or as T_1 approaches T_0, one is the upper limit on the slope. This is also verified in the Handbook H108 tables.

The previous discussion of sampling indicates that as the test ratio approaches one, the test should be difficult to pass. Later discussion in this paper will show this to be true. In general, as the test ratio increases, the decision line slope increases and, if changes in the h intercepts (see Figure 6) do not offset the slope changes, the test will take longer and be more difficult to pass.



T_1/T_0	s/T_0
0.25	0.462
0.33	0.549
0.50	0.693
0.67	0.811
1.0	1.0 (theoretical)

Fig. 7

s/T_0 Variation with T_1/T_0

<u>Variation of Intercepts h_0/T_0 and h_1/T_0 with T_1 and T_0</u>

The accept H_0 intercept on the $V(t)/T_0$ test time axis is h_0/T_0. The reject intercept is h_1/T_0. Both are functions of all four test parameters. Eqs 7 and 8 are derived from Eqs 1, 2, 3, and 5.

$$h_1/T_0 = \frac{-\ln\left(\frac{1-\beta}{\alpha}\right)}{(T_0/T_1) - 1} \quad (7), \qquad h_0/T_0 = \frac{-\ln\left(\frac{\beta}{1-\alpha}\right)}{(T_0/T_1) - 1} \quad (8)$$

A good feature of the normalised decision relationship is the use of the same h intercept values for many absolute values of T_1 and T_0. Whether the values be 30 and 60 or 750 and 1500 the same normalised intercept applies. Assuming α and β constant, at low test ratio T_1/T_0 the denominator of the Eqs 7 and 8 is large and both h intercepts are small. This is certainly reasonable. As the continue-test width narrows a decision is expected sooner. This is indicative of a test with low discrimination as will be discussed later. A low test ratio indicates T_1 much less than T_0 which is distant from the nearly perfect O.C. curve case discussed earlier.

<u>Variations of Intercepts h_0/T_0 and h_1/T_0 with α and β</u>

The test ratio is assumed constant in this section. The reject intercept (Eq 7) is a function of A or $(1-\beta)/\alpha$. As α , the producer's risk, decreases, A increases, lnA increases in magnitude since A is greater than one, and the reject line becomes more negative or the continue-test region is widened on the reject side. This seems sensible

since α is the probability of rejecting acceptable or good items.

As β decreases, lnA again increases and the reject line is lowered and the no-decision zone increased.

For the acceptance intercept (Eq 8), as α decreases, $\beta/(1-\alpha)$ (or B) decreases, lnB becomes more negative since B is less than one, and the accept intercept becomes more positive. Therefore, as α decreases, the accept line moves away from the r s/T_0 line (see Figure 6), the zone of no decision is widened on the high or accept side, and there is a smaller chance that a particular event point will fall in the accept area.

As β decreases, B decreases. Again, the accept intercept becomes more positive or larger.

The magnitudes and directions of changes in the intercepts might be determined and analysed by taking partial derivatives of Eqs 7 and 8 as in the following examples. Such an analysis was not conducted for this paper. For brevity, $(T_0/T_1) - 1$ is termed w.

$$\left.\frac{\partial (h_0/T_0)}{\partial \beta}\right|_{\alpha} = \frac{\partial \left[\dfrac{-\ln\left(\dfrac{\beta}{1-\alpha}\right)}{w}\right]}{\partial \beta} = \frac{-1}{w}\,\frac{\partial\left[\ln \beta - \ln (1-\alpha)\right]}{\partial \beta} = \frac{-1}{\beta\, w} \tag{9}$$

$$\left.\frac{\partial (h_1/T_0)}{\partial \beta}\right|_{\alpha} = \frac{\partial \left[\dfrac{-\ln\left(\dfrac{1-\beta}{\alpha}\right)}{w}\right]}{\partial \beta} = \frac{-1}{w}\,\frac{\partial\left[\ln (1-\beta) - \ln \alpha\right]}{\partial \beta} = \frac{-1}{w}\,\frac{(-1)}{(1-\beta)} = \frac{1}{(1-\beta)\, w} \tag{10}$$

__Variations of O.C. Curve__

Appendix C describes a method of calculating the complete O.C. curve for any particular test. A simple mathematical relationship for L(T), the probability of accepting H_0, in terms of α, β, and T_1/T_0 is not available except for isolated points. The simplest approach is to use the data and O.C. curves of Handbook H108 (Ref 18:2.9-2.20,2.63-2.64). This approach was used in the following sections of this chapter.

__Variation of O.C. with T_1/T_0.__ As the test ratio increases, the slope of the O.C. curve increases and the test becomes more discriminating. Figure 8 gives sketches of the O.C. curves for test plans taken from Handbook H108. The C codes are for test plans with α of 0.1 and β of 0.1. The D codes are for test plans with α of 0.25 and β of 0.1 (Ref 18:2.2). If α, β, and T_0 are constant, as T_1/T_0 increases, the probability of accepting a lot with a particular T/T_0 value decreases if the T is below T_0. Figure 8 illustrates this. For example, if T/T_0 is 0.6, as T_1/T_0 increases (code C-10 to C-15) the L(T) decreases. If the T true is greater than T_0, the L(T) increases as T_1/T_0 increases.

__Variations of O.C. with T_1 Constant.__ Figure 9 compares the O.C. curves of two tests both with α and β risks of 0.1 but with different test ratios. The curves of Figure 8 were multiplied by the appropriate T_0/T_1 ratio to give the abscissa normalized in terms of T_1 rather than T_0. For any particular test this makes no difference. The absolute T will be the same regardless of how it is normalized. [illegible]

Code	T_1/T_0
C-10	0.438
C-15	0.666
D-5	0.421
D-12	0.650

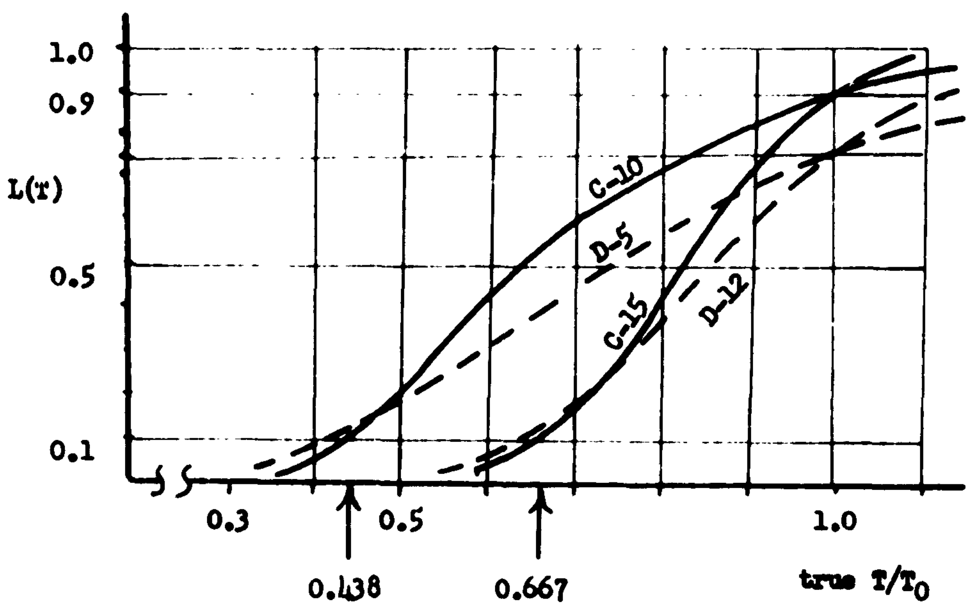


Fig. 8

Variation in O.C. with T_1/T_0

(From Ref 18:2.15-2.19)

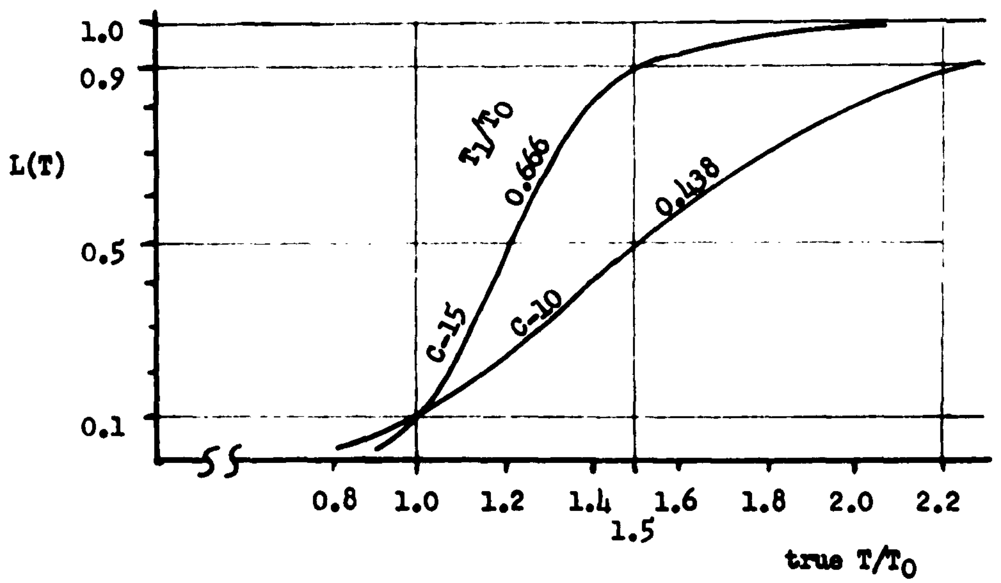


Fig. 9

O.C. Normalised to T_1

(From Fig. 8)

figure indicates, for a set β risk, as the T_1/T_0 ratio increases, the slope increases and the producer can obtain an α of say 0.1 at a lower T true than with a lower T ratio test. Changing the normalization merely slides the curve in Figure 8 across the abscissa. If the consumer has set some definite T_1 and β which must be held, such a change in normalization could be useful in determining what the contractor must meet for the various T ratio tests to be considered.

__Variation of O.C. with α.__ Figures 10 and 11 illustrate the changes in probability of acceptance when α is increased to a level as high as 0.25 with β constant. The increase in α reduces the probability of acceptance for all T/T_0 values greater than T_1/T_0 for a test with the same T_1/T_0 test ratio. As α increases with constant β and T_1/T_0 ratio, the O.C. curves tend to pivot about the T_1/T_0, β point and the probability of accepting lots below T_1/T_0 increases slightly. Codes C-10 and D-5 in Figure 8 also are examples. Therefore, a very high α or producer's risk does not help the consumer in all respects.

It seems doubtful that a producer would like to accept an α risk as high as 0.25, but if his expected T were much greater than T_0 he might in order to reduce the test time.

__Variation of O.C. with β.__ The curves are not shown, but by analogy to the discussion for α variations above, when the β for a particular test is increased, the O.C. curve, for example of C-10 in Figure 8, would pivot about the L(T) equal 0.9, T/T_0 equal 1 point and the L(T) for T less than T_0 would increase. For T greater than T_0,

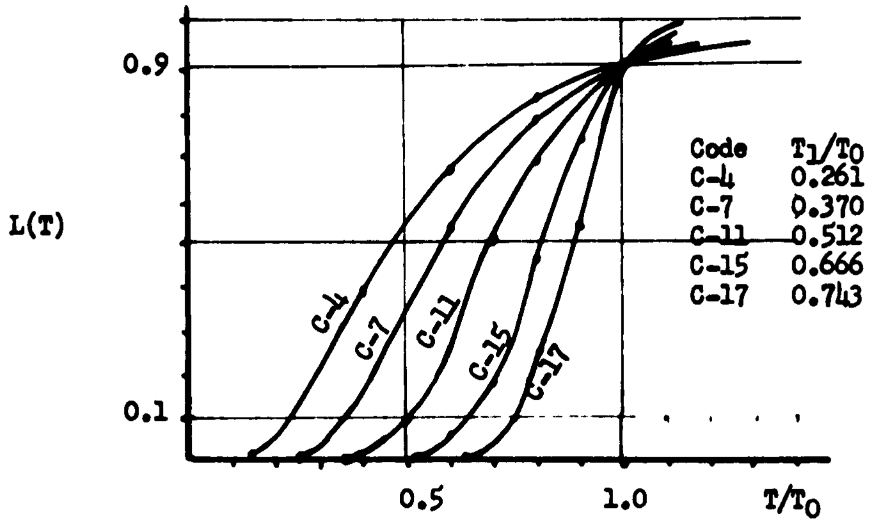


Fig. 10

Variation of O.C. with T_1/T_0, α of 0.1
(Data from Ref 18:2.2,2.13)

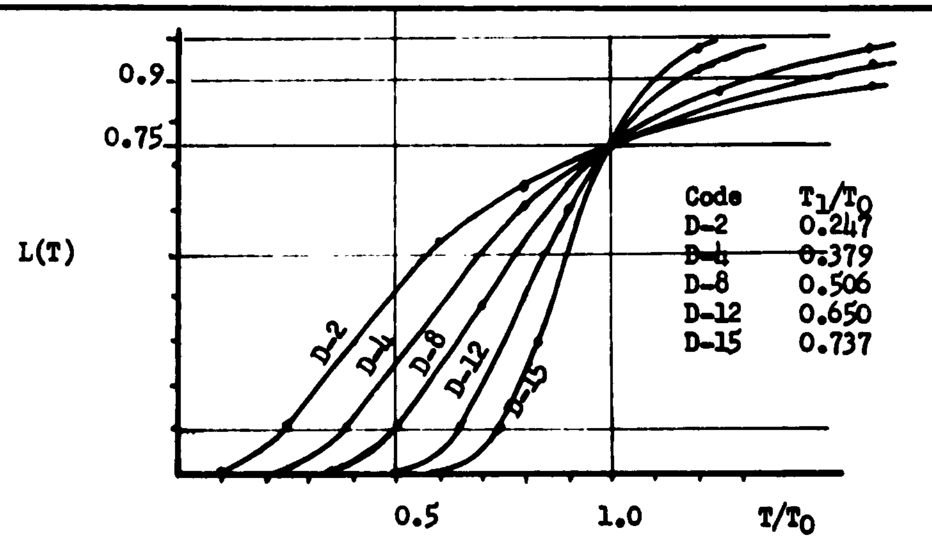


Fig. 11

Variation of O.C. with T_1/T_0, α of 0.25
(Data from Ref 18:2.2,2.17)

the probability of acceptance, L(T), would decrease. Therefore, the producer has an interest in holding β within reasonable limits when his product is very good.

The cross effects of a high β increasing the effective α above T_0 and of a high α increasing the effective β below T_1 are diminished, but they do exist and can help the desire for reasonable compromises when both sides are willing to admit the cross effects exist.

Variation of AFN Average Failure Number

A characteristic of any SPR test that is of major interest is the expected or average failure number before a decision is reached, AFN. In sequential tests in which the unit of test is an item not a failure, such as tests of fraction defective, this is the ASN or average sample number. As mentioned previously, in mean time tests each specimen or unit of test can be considered one failure. Appendix D gives the equations that can be used to develop the AFN for the tests and for various T/T_0 values. The relationships involve L(T) which is derived from a pair of parametric equations (See Appendix C). This creates difficulty in developing a simple mathematical relationship for the AFN and the related expected test time in terms of the four test-defining parameters, α , β, T_1, and T_0 (or T_1/T_0). Handbook H108 (Ref 18:2.63-2.65) gives values of AFN (termed $E_0(r)$ in that reference) for many SPR tests. These are recommended for use.

Variation of AFN with T_1/T_0. Figures 12 and 13 indicate that for α and β constant, as the T_1/T_0 ratio increases, for any particular true T/T_0, the AFN increases. Or, as the discrimination of the test goes up, the expected failures go up.

Variation of AFN with α and β. If T_1/T_0 is essentially constant (for examples C-15 and D-12 in Figures 12 and 13) and α decreases (Figure 13 to Figure 12) the AFN will increase. Higher assurance that T_0 quality items will not be rejected requires more testing. A decrease in β is not so obvious from the figures but, as explained earlier, and as suggested by AMCP 74-1 (Ref 16:3), an increase in β is equivalent to a lower T_1/T_0 ratio test at the same β risk. Therefore, a decrease in β can be represented by an increase in the T_1/T_0 curve used. Again, the AFN increases for a greater assurance that items as bad as T_1 will not be accepted or for a decrease in β.

Maximum AFN. It should be noted that the maximum AFN for any one test occurs when T lies between T_1 and T_0 in the vicinity of s (see Eqs 5,6). Therefore, if the true T of the items can reasonably be estimated, the test selected should not place that T near s. This would be deliberately aiming for the longest test possible.

It will benefit no one to set T_1 above the expected T. The consumer will seldom benefit from rejection of the bulk of the product submitted. If the T_1 needed were actually higher than the T possible to produce and T_1 could not be reduced, then screening inspection or some other form of 100 % sorting should be considered. A sequential test cannot upgrade the quality submitted any more than any other test

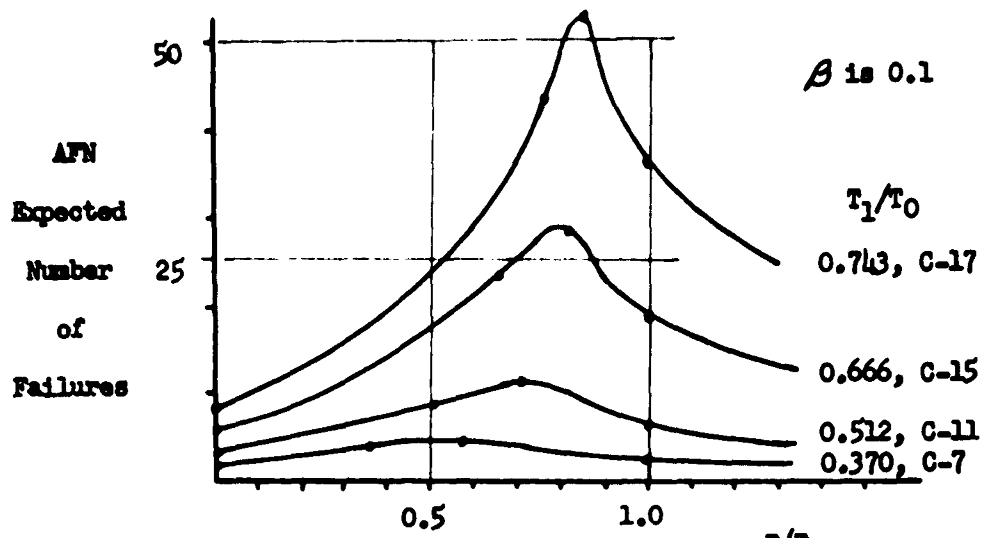


Fig. 12

Variation of AFN with T_1/T_0, α of 0.1

(Data from Ref 18:2.69)

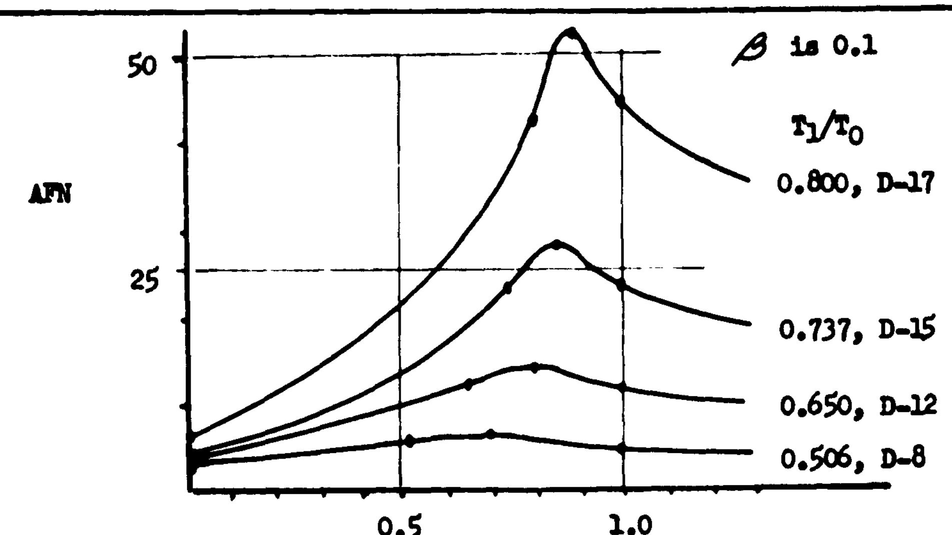


Fig. 13

Variation of AFN with T_1/T_0, α of 0.25

(Data from Ref 18:2.64)

can. It can give information at a minimum cost, and it can provide incentive to the producer not to reduce quality excessively through over-diligent cost-reduction programs.

In general, the T expected should be above T_0 or at most slightly below it depending on the preference of the producer to take or avoid risks and on the consumer's desire for a short or long test. A long test may provide more information on the equipment, but it will probably delay delivery.

Small AFN. As indicated by Figures 12 and 13 a smaller AFN occurs when the equipment is very poor or very good. This agrees with the general concept of a sequential test. If the first 17 walnuts sampled from one bag were good, the probability that the whole bag was good would certainly be much higher than if 17 of 25 were good. If the equipment shows very good or very bad T the test can be cut short. Thus, a faster test occurs when the true T is outside the T_1 to T_0 band.

Variations in Expected Test Time

Appendix D indicates that the expected waiting time for a decision is a direct function of the AFN and is quite direct when failed items are replaced or repaired and returned to the test soon. It is difficult to state exactly how much off-test time in repair is too much, but a down time of 5 % or less of the operating time should not cause an error more than the other estimates cause. This figure is admittedly subjective and should be investigated on individual tests if it appears

troublesome.

The same comments that applied to the AFN apply to $V(t)/T_0$, the sum of the expected waiting time in multiples of T_0. As T_1/T_0 increases, the waiting time increases. As α or β risks decrease, the waiting time increases - Figures 14 and 15, as could be expected, are similar in form to those for AFN versus T/T_0 except the influence of increasing T tends to reduce the drop off of $V(t)/T_0$ above T_0. The reduction in test time is still appreciable, however. The maximum expected waiting or test time is again when T true is between T_1 and T_0 and near s.

Incentive for High T. The reduction in test time and therefore in test costs should act as a strong incentive on the producer to produce items with T greater than T_0. Of course, it also supplies incentive for him to agree to the lowest T_0 possible. The reduction for various T greater than T_0 was not calculated for this paper. AMCP 74-1 (Ref 16:11) indicates that for AGREE Task Group 3 tests, where T_1/T_0 is 0.667, if the true T is twice T_0, the expected test time is on the order of one third the expected test time when true T is T_0. Figure 14 indicates that the savings will be especially large if tests of high discrimination or high T_1/T_0 ratios are used. Figure 15 indicates that the savings will be less at high α tests.

Variations of Test Character with Sample Size n

If the distribution of failures is actually exponentially distributed the number of equipments, n, put on test should have no effect

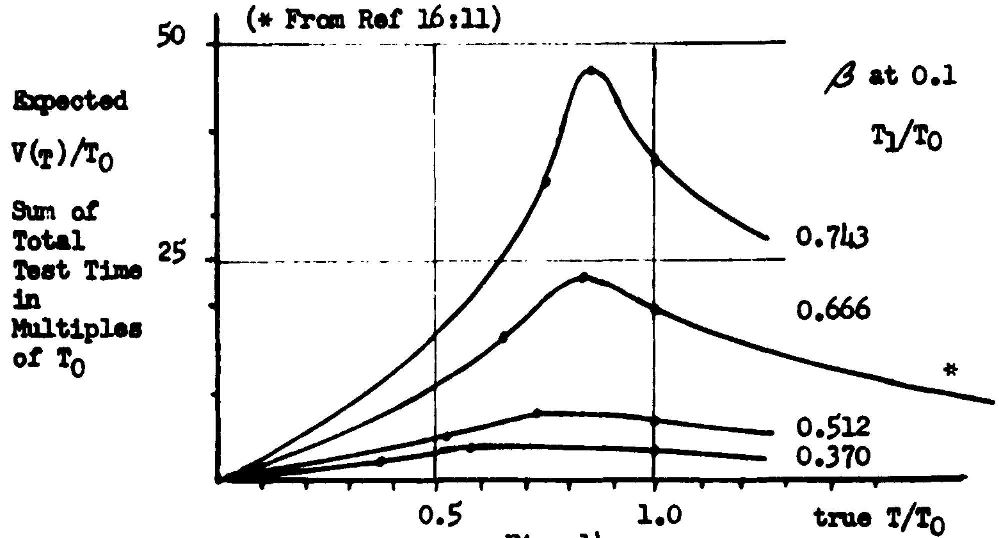


Fig. 14

Variation of Test Time with T_1/T_0, α at 0.1
(From Ref 18:2.64 data and Table II)

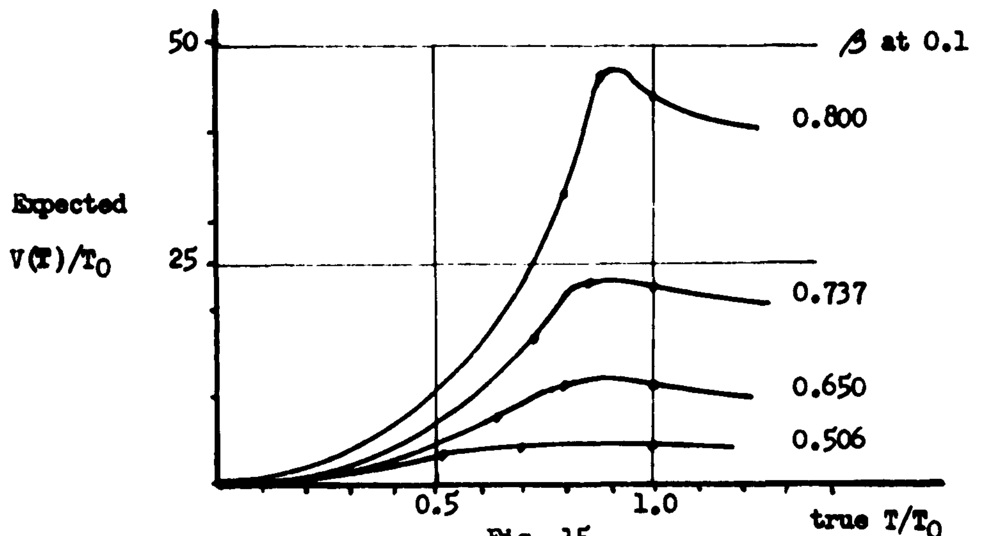


Fig. 15

Variation of Test Time with T_1/T_0, α at 0.25
(From Ref 18:2.64 data)

on the test other than to reduce the calendar waiting time if n is large. For a particular lot, the sum of the test or waiting time on all equipments should be the same regardless of the size of n. The tests considered in this paper are primarily concerned with small n samples and small lots. The problems of sampling as covered by the production test procedures of AGREE Task Group #3 (Ref 13:167-174,181-190) are not covered Currently, on relatively small lots the only reasonable approach seems to be to assume that the few items that can be placed on test truly represent the lot while adding whatever requirements seem practical to increase the probability that the assumption is true.

Distribution not Exponential

There is, of course, the very real risk that the distribution is not exponential. This section is the only one in this paper considering a non-exponential distribution. The distribution may have an increasing or a decreasing failure rate with time such as the Weibull distributions described by Technical Reports TR-3, -4, and -6 (Refs 20, 21, 22). A decreasing failure rate could be found for some parts, but in equipments it would probably indicate inadequate burn-in or debugging by the producer before the items were submitted for test. If the failure rate were truly decreasing with time, a large n on test would mean that the T seen by the test would tend to be short. Fewer items on test would give the items time to get out to the longer T life period.

The more probable variation from exponential will be a failure rate increasing with time. If n were large, the sum of test hours for all n items would accumulate rapidly and the test would in effect see a low failure rate or high T. The probability of accepting equipment that would seem poor in later life would be high. Low n would stretch the test time on each item so the higher failure rate period could be seen. Usually, there will be far from an abundance of equipments available for reliability test. A small n will at least tend to reduce the increasing failure rate problem, but a small n can also reduce the probability that the sample truly represents the total lot being presented for acceptance.

Since large amounts of data are usually needed to determine with any high degree of confidence the true distribution of a population, it will probably be difficult to establish tests for new equipments with any high degree of assurance that the distribution assumptions are sound. When no data are available indicating otherwise, the exponential distribution seems a reasonable assumption. It could be said that it takes the middle ground if the distribution is completely unknown, it greatly simplifies calculations, much work using the exponential assumption that has been done in operations research and telephone systems, and as Epstein points out it allows application of the well developed Poisson process (Ref 7:2). Epstein's work on tests to determine if assumption of the exponential distribution is valid is an excellent reference for methods to be applied

during or after a test (Ref 10).

Truncation

In any SPR test there exists the possibility that the test might terminate only after a long time regardless of the α, β, T_1/T_0 parameters. Usually, the probability that an SPR test will terminate before an equivalent conventional test (same α and β) is on the order of 80 %, but a particular test sequence could wander about in the no decision zone. It is suggested that the procedures of Handbook H108 be used in specifying truncation points which cut-off or stop the test at some combination of test time and failures when the time has grossly exceeded that expected. The convenience of the procedure is one more advantage in using H108 (Ref 18:2.58,2.60,2.62). It is believed that a procedure to simplify calculations of α and β risks at much earlier truncation points would be useful, but this was not attempted for this paper.

V. Test Costs and Selection Procedure

Some relationship between test cost or time and the T_1/T_0 test ratio could be very helpful in selecting a good test plan for each particular problem (the term optimum test is avoided since an optimum test based on rough estimates seems unlikely). The SPR test plan data of Handbook H108 (Ref 18) seems to have been based on the methods described in this paper, so data from that handbook were used as a starting point.

Relationship of Test Time to T_1/T_0

In attempting to find a relationship between the expected test time and the test ratio it seems reasonable to assume that few contractors will be pleased about an α risk greater than 0.1. As previously discussed, even if the consumer estimates a T_1 point at β other than 0.1, the test can be described by some other T_1 or T_1/T_0 at β equal to 0.1. Therefore, the Code C test plans of Handbook H108 were used. These plans and the associated O.C. curves set α at 0.1 at T_0 and β at 0.1 at T_1 (Ref 18:2.13-2.16,2.64).

The AFN values for T equal to T_1, s, and T_0 were read from Handbook H108 for various test plans. That data and the data resulting from the procedure described below appear in Appendix E. Using the relationship of $V(t)/T_0$ to AFN described in Appendix D, the time $V(t)/T_0$ was calculated. That average accumulated test time in multiples

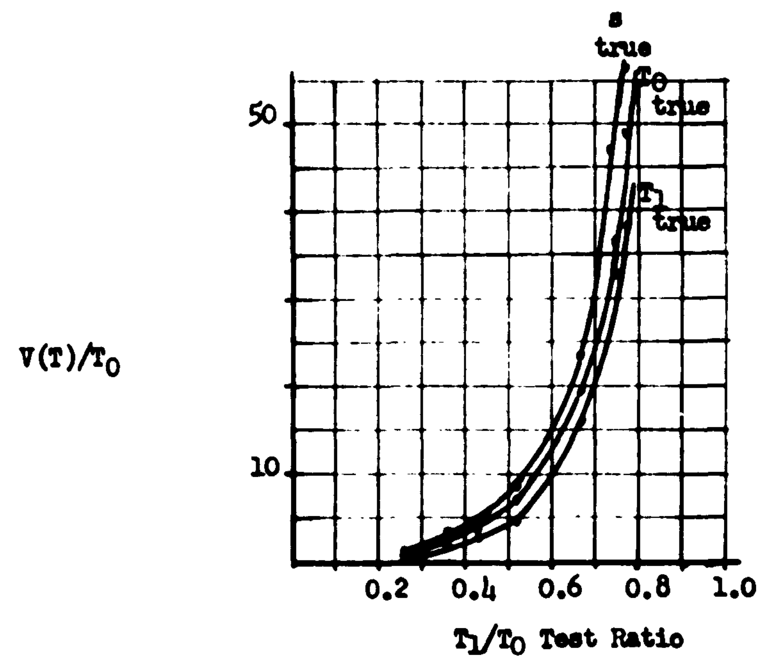
of T_0 versus the T_1/T_0 is sketched in Figure 16. As could be noted from Figures 14 and 15, if the T true is equal to T_1, the expected test time is shortest of the three cases considered. If true T is equal to s, the slope factor, the test time is longest.

Regression of Test Time on T_1/T_0

The shape of Figure 16 indicates an exponential relationship. The regression method used is described in Appendix E. Probably due to the $T_0/T_1 - 1$ terms in the inequalities governing the SPR test, the relationship does not seem exponential over the entire range. However, the calculations indicate that for a T_1/T_0 range of approximately 0.26 to 0.72 the regression equation below holds within plus or minus 7 %.

$$V(T_0)/T_0 \cong e^{(7(T_1/T_0) - 1.7)} \tag{11}$$

This relationship gives the expected time to reach a decision in multiples of T_0 when the true T is T_0 as a function of the test ratio T_1/T_0. T_0 seemed a reasonable compromise T to use as an approximation of the T to expect. Using the curve for T equal to s (s true in Figure 16) seemed pessimistic. It could be hoped that the average test did not face this high test time condition. If, on the average, T_1 is true, both the producer and the consumer will be in deeper trouble than long test time. The producer will experience a rejection rate of $100(1 - \beta)$ % of production. The consumer will receive poor items primarily and not many of even these. T_0 is the target



Code C Plans Ref 18,
α and β of 0.1

Fig. 16

Sketch of Test Time Versus Test T_1/T_0 Ratio

(Data from Appendix E,
Table III)

point at which the producer will probably aim. Figure 17 shows the regression plot of Eq 11.

From the regression it seems that some relatively simple relationship can be obtained for the expected test time as a function of the test ratio. If this relationship could be combined with the variables of the physical test procedure such as number of hours per day, shifts per day, days per month, test facilities per unit, and the several others that can change the cost of a test and the calendar time required for a decision, then a test cost as a function of the test (T_1/T_0) could be developed. Such a relationship was not developed for this paper, but one seems feasible.

Costs

The Naval Ordnance Test Station Manual indicates that one valid relationship for determining sample size is

$$\begin{aligned} C_T = \; & \beta \text{ (Cost of Accepting } H_0/H_1 \text{ true)} \\ & + \alpha \text{ (Cost of Rejecting } H_0/H_0 \text{ true)} \qquad (12) \\ & + \text{Cost of Test or Experiment} \qquad \text{(Ref 3:17)} \end{aligned}$$

or

$$C_T = \beta C_B + \alpha C_A + C_e \qquad (13)$$

where

C_T is total cost

α is producer's risk

β is consumer's risk

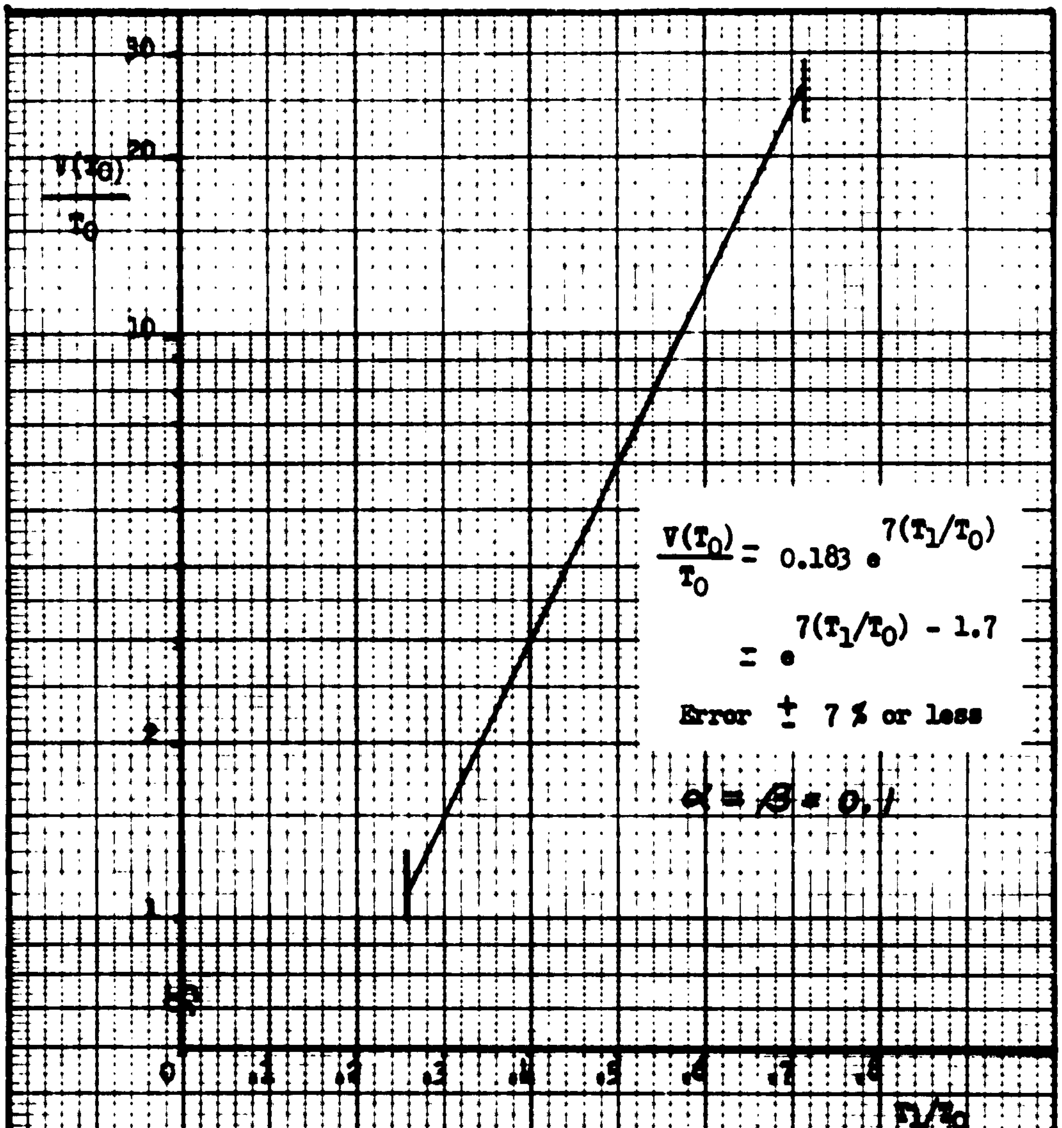


Fig. 17

Relationship of $V(T_0)/T_0$ to

T_1/T_0 Temp Ratio

(Data from Appendix I and Table III)

β C_B is the first term of Eq 12

α C_A is the second term of Eq 12

C_e is test cost

If a relationship for the above quantities in terms of T_1/T_0 could be developed the total cost, C_T, could be minimized with respect to T_1/T_0 rather than minimized with respect to n, the sample size. In this section as in most of this paper, the sample size available for reliability test is considered essentially constant. Whatever units that can be made available (which will probably be few) will be used as n.

<u>C_B Cost</u>. A relationship for C_B of Eq 13 in terms of T_1/T_0 is desired. It is assumed the problem concerns a checkout equipment used in what could be considered a critical application supporting a combat vehicle. The probability that the equipment will not fail through its mission or use time, t, is

$$P_c = P(\text{checkout not fail in } t) = e^{-t/T} \tag{14}$$

The probability that the combat vehicle will fail due to a checkout failure is defined as

$$Z = P(\text{vehicle fail/checkout failed}) = \frac{\text{Number of vehicle failures due to checkout failures}}{\text{Number of checkout failures}} \tag{15}$$

The probability that the combat vehicle will fail its mission due to checkout failure (or error) is

$$P_f = (1-P_c)Z = P(\text{vehicle fail due to checkout fail}) \tag{16}$$

This is based on the assumption that the designs of the checkout equipment and the vehicle are adequate and compatible - that is when the checkout behaves in what is defined as proper operation, the vehicle is properly checked. This, of course, assumes away a difficult task. If C_f is defined as the cost of a lost mission, then the expected lost mission cost per mission due to checkout failure is

$$C_m = C_f P_f = C_f(1 - P_c)Z \tag{17}$$

If the distribution of checkout failures can be assumed exponential, from Eq 14

$$C_m = C_f Z(1 - e^{-t/T}) \tag{18}$$

If an estimate of the cost when T equals T_1 is desired, and if the exponent is normalised to T_0, Eq 18 becomes

$$C_B = C_f Z\left(1 - e^{-\left[\frac{t/T_0}{T_1/T_0}\right]}\right) \tag{1}$$

where

C_B is expected lost mission cost due to checkout failures when T_1 is true or the cost of accepting H_0/H_1 true

C_f is cost of a lost mission

Z is defined by Eq 15

t is operating time of checkout equipment

T_0 is the high target point as earlier

T_1 is the assumed actual T of the checkout equipment

As T_1/T_0 decreases, C_B will increase which is to be expected. If T_0 is held constant, as the T_1 decreases more failures are expected. Eq 19 should be appropriate for the first terms of Eqs 12 and 13.

Repair and maintenance costs would be added in Eq 19 if appreciable, but for simplicity here, it is assumed a lost mission cost is much greater.

C_A Costs. The cost of rejecting H_0 given H_0 is true could be considered as constant with reference to T_1/T_0. These costs of rejecting the lot when T_0 is the true T will involve rework costs, resubmission costs, and whatever value the producer puts on any reputation loss. The consumer loses the availability of the actually acceptable equipment which might be badly needed. These costs will essentially not vary with the T_1/T_0 test ratio.

If Eq 12 could be developed in terms of T_1/T_0, the conventional approach of taking the derivative with respect to T_1/T_0 and equating to zero should lead to the minimum cost. This procedure would in effect equate marginal cost to marginal gain. The C_A costs would drop out of the relationship since they would not vary with T_1/T_0. These costs would change the magnitude of the minimum cost point, but they would not change its position with respect to T_1/T_0, which is the parameter desired.

C_e Test Costs. The test cost will be function of the following test conditions: the complexity of the measurements to be made, the environmental facilities required, the measurement and test equipment

required, the personnel skills required, the penalty costs involved by second and third shift operation and six and seven day operation, the power required, the number of equipments on test, the length of test, and other factors. For any particular n, it should be possible to estimate C_e as a function of nt for various values of the above factors controlling test cost. Since n can be considered essentially constant for one particular problem, the test time t will be the major variable factor in the nt value. There is no reason to expect that all C_e functions would be linear, but even linear approximations made from a few calculated cost points should indicate the general trends of cost change.

Minimum Cost. The $V(T_O)/T_O$ of Eq 11 is nt/T_O. Therefore, if the proper algebraic manipulations are applied to Eq 11 and a cost function as discussed in the preceding paragraph, it should be possible to obtain an expression for C_e as a function of nt, T_O, and T_1/T_O. A graphical approach using Figure 17 and plots of the cost functions with nt as the independent variable could lead to plots of the summation of expected costs. The combination of C_B and C_e as in Eq 13 should indicate the minimum cost point and the related T_1/T_O indicated as the best test ratio.

The algebraic work suggested above was not completed for this paper. The approach seems feasible, however, and if developed should

yield at least a good starting point in selecting the T_1/T_0 ratio for a particular test.

<u>Procedure for Selecting an SPR Test</u>

The procedure above certainly is far from a final answer on a method to select a test. Appendix J contains a suggested approach to selecting a reasonable T_1/T_0 test for a particular job. The suggestion is primarily to use brute force in trying a T_1/T_0, reviewing the risk considerations of Appendix F, estimating the costs and time, then trying again if the result is not satisfactory. Hardly a sophisticated procedure, but it seems to be the one recommended by much of the literature on the subject. Again, it is recommended that H108 be used at least for guidance when appropriate.

VI. Conclusion

Results

A step-by-step method for a guide in selecting a reasonable SPR (sequential probability ratio) test was desired. Appendix J gives such a guide. As previously mentioned, it is hardly a sophisticated method, but it should be helpful in obtaining tests nearer optimum than pure guesses. The flow chart of that appendix references several of the other appendices and tables and Figure 17 of this paper in the hope of making best use of the effort invested.

Appendix F lists some of the points to consider in choosing acceptable risks for a test.

Figure 5 and similar curves that could be obtained from the methods described in Appendix D using Handbook H108 should supply information to show project engineers what an SPR test could do to reduce waiting time before a test decision.

Table I, which is included here as a summary of some of the previous work, gives an easily obtained reminder of the effect of changing any of the SPR test parameters. It could be useful in making decisions on how to change a test plan that is unacceptable.

In general then, the results of this paper consist of a series of charts and lists intended to help in selecting and in modifying SPR tests.

Table I
Summary of Test Parameter Effects

parameter	change parameter ← all others constant	Effect on AFN V(t) test length	Prob. Accept Marginal Units $T<T_1$	Prob. Reject Good Units $T>T_0$	accept line h_0 relative to rs	reject line h_1 relative to rs	slope s. and sr line
T_1/T_0	decrease	decrease	higher	higher	toward	toward	decrease, toward reject
T_1/T_0*	increase	increase	lower	lower	away from	away from	increase, toward accept
α	increase	decrease	higher slightly	higher	toward somewhat	toward	independent of α
α *	decrease	increase	lower slightly	lower	away from somewhat	away from	"
β	increase	decrease	higher	higher slightly	toward	toward	independent of β
β *	decrease	increase	lower	lower slightly	away from	away from	"
n	increase	decrease	no influence on test strength if exponential is true. Increases P (accept poor items) if hazard rate increases.				
n.	decrease	increase	no influence if exponential true. Decreases P (reject good items) if hazard rate decreases.				
T true	between T_1 and T_0	longest	no influence on test plan and decision lines.				

* higher discrimination or ability to separate good and bad lots.

Conclusions

SPR tests will on the average yield shorter waiting times for a decision than conventional tests. The time savings vary with the parameters of the test, of course, but the savings can be on the order of 50 %.

Handbook H108, Sampling Procedures and Tables for Life and Reliability Testing (Based on the Exponential Distribution), is an excellent reference for various test plans of the SPR (sequential probability ratio) tests discussed in this paper. If an existing military specification or other standard does not exist giving more test plan and procedure details, for a particular test selected, that handbook should be used whenever possible.

Figure 17 indicates that there is some relatively simple exponential relationship within a useful range of T_1/T_0 between $V(T_0)/T_0$ and T_1/T_0 which are, respectively, the sum of the expected test time in multiples of T_0 when the true T (MTBF) is T_0 and the test ratio of the lower and the upper specified limits of T for the particular plan. The discussion in Chapter V indicates that if the proper mathematical relationship could be derived, selecting a good test plan could be greatly simplified, but that a graphical approximation may be practicable in some problems.

In general, SPR tests should be considered for reliability tests of most equipment. Simple equipments for which adequate design can be completely specified are another question. Commonly used life tests involving preassigned times for termination such as those described in the front sections of Handbook H108 and by Epstein (Refs 18, 9, 11) are certainly valid tests and are appropriate in many instances, but SPR tests should always be at least considered. Basovsky is a convenient

reference for a description of the preassigned time tests (Ref 1:238). ANCP 74-1, Quality Control Reliability Evaluation Procedures for Pilot-Production and Production as Recommended by Task Group No. 3 of AGREE, is an excellent reference for assistance in understanding and in using AGREE procedures.

Recommended Future Work

Investigate the relationships between the $V(t)/T_0$ expected test times and the T_1/T_0 test ratios with the intention of developing the mathematical relationships for minimum test cost and expected loss as functions of the test ratio. Appendix E, Eq 11, and Figure 17 indicate that such a relationship should be feasible. It would be a very useful tool if available.

Study production multiple sampling procedures for reliability tests with the intention of outlining the procedure to be followed in establishing various plans for various risks and production rates. Reference is made to the production sampling procedure recommended by AGREE Task Group 3 (Ref 13:167).

Investigate the possibility of and the practicality of developing a series of test plans similar to Handbook H108 for distributions with increasing hazard rates with time. The Weibull distribution is, of course, an example.

Study truncation of SPR tests with the intention of outlining a procedure for establishing the risks incurred by various truncation times.

Bibliography

1. Basovsky, I. Reliability Theory and Practice. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1961.

2. Carhart, R. R., A Survey of the Current Status of the Electronic Reliability Problem. (U) RM-1131. Santa Monica, Calif: Rand Corp, 14 Aug. 53. AD80637

3. Crow, E. L., et al. Statistics Manual. NavORD Report 3369. China Lake, Calif.: U. S. Naval Ordnance Test Station, 1955.

4. Duncan, A. J. Quality Control and Industrial Statistics. Homewood, Ill.: Richard D. Irwin, Inc., 1959.

5. Epstein, B., and Sobel, M. Sequential Life Tests in the Exponential Case. Detroit, Mich.: Wayne University, 31 Jany. 54. AD26 141.

6. Epstein, B. Statistical Developments in Life Testing. Wayne State University: Detroit, Mich., 1 June 57. AD155-353

7. -----, The Exponential Distribution and Its Role in Life Testing. Detroit 2, Mich.: Wayne State University, 1 May 58. AD158-605

8. -----. Statistical Techniques in Life Testing, Chapter II, Testing of Hypotheses. Detroit 2, Mich.: Wayne State University, 1 Oct. 58. AD211-457.

9. -----. Problems of Estimation. Detroit 2, Mich.: Wayne State University, 15 Jany. 59. AD211-458.

10. -----. Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential. Detroit 2, Mich.: Wayne State University, 2 April 59. AD215-402.

11. -----. Tolerance Limits Based on Life Test Data Taken From an Exponential Distribution. Detroit 2, Mich.: Wayne State University, 1 Feb. 59. AD211-456.

12. Lloyd, D.K., and Lipow, M. Reliability: Management, Method, and Mathematics. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.

13. Reliability of Military Electronic Equipment. Advisory Group on Reliability of Electronic Equipment. Washington, D. C.: OASD(R&E), 4 June 57.

14. Sequential Analysis of Statistical Data: Applications. Statistical Research Group. Columbia University. New York: Columbia University Press. 1945.

15. Wald, A. Sequential Analysis. New York: John Wiley & Sons. 1947.

16. AMCP 74-1. Quality Control Reliability Evaluation Procedures for Pilot-Production and Production, as Recommended by Task Group No. 3 of AGREE. Air Materiel Command Pamphlet. Wright-Patterson AFB, Ohio: Hq. AMC. 31 March 1959.

17. H 105. Administration of Sampling Procedures for Acceptance Inspection. Wash., D. C.: OASD(S&L). 27 April 1954.

18. H 108. Sampling Procedures and Tables for Life and Reliability Testing (Based on the Exponential Distribution). Wash., D. C.: OASD(S&L). 29 April 1960.

19. MIL-R-26667A(USAF) (2). Reliability and Longevity Requirements, Electronic Equipment, General Specification For. Wright-Patterson AFB, Ohio. 1 July 63.

20. TR 3. Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (Mean Life Criterion). Wash., D. C.: OASD(I&L). 30 Sept. 61.

21. TR 4. Title as above except...(Hazard Rate Criterion). Wash., D. C.: OASD(I&L). 28 Feb. 62.

22. TR 6. Title as above except...(Reliable Life Criterion). Wash., D. C.: OASD(I&L). 15 Feb. 63.

23. WD-25477. vol. 5. Naval Weapons Data, Reliability Engineering, Reliability Testing. Wash., D. C.: Bureau of Naval Weapons, 15 Dec. 62.

Note: All the documents by Professor Epstein were sponsored at least in part by the Office of Naval Research.

Appendix A

SPR Test for Exponential or Poisson Case

If the underlying distribution of failure of a population is governed by a constant instantaneous failure rate or hazard rate, the probability density function will be of the form

$$f(t,T) = \frac{1}{T}e^{-t/T} \; ; \; t>0, \quad T>0 \tag{20}$$

where T is the mean time between failures and t is the mission or operating time. The cumulative distribution from 0 to t is the unreliability or the probability of one or more failures in time t. This leads to the well known reliability relationship or probability of no failures of

$$R = e^{-t/T} \tag{21}$$

which is also the probability of no failures of a Poisson distribution with MTBF of T. The exponential then actually describes a Poisson process.

If there are n units on test in a Poisson process the failure rate will be in effect nt/T. Therefore, the probability of exactly r failures in time t is

$$P(r;n/T) = \frac{(nt/T)^r e^{-nt/T}}{r!} \tag{22}$$

Therefore, the SPR or liklihood ratio for any observed r is

$$\frac{P(\text{getting sample got}/H_1 \text{ true})}{P(\text{getting sample got}/H_0 \text{ true})} = \frac{P_1}{P_0} = \frac{P(r;nt/T_1)}{P(r;nt/T_0)} =$$

$$= \frac{\frac{1}{r!}\, e^{-(nt/T_1)}\, (nt/T_1)^r}{\frac{1}{r!}\, e^{-(nt/T_0)}\, (nt/T_0)^r} = e^{-(nt/T_1)+(nt/T_0)}\, (nt/T_1)^r\, (nt/T_0)^{-r} =$$

$$= (T_0/T_1)^r\, e^{-\left(\frac{1}{T_1} - \frac{1}{T_0}\right) nt} \qquad (23)$$

where test is for $\quad H_0: \ T = T_0 \quad (T \geq T_0)$

$\qquad H_1: \ T = T_1 \quad (T \leq T_1), \quad (T_1 < T_0) \qquad (4)$

This, as it should, agrees with Epstein and Sobel (Ref5:2) and Basovsky (Ref1:258). It should be noted that nt is the total test time for all equipments on test. For a replacement test, this is obviously true. For a nonreplacement test, the sum of the various test times on the different equipments is used. This total test time is termed V(t) in this report which agrees with much of the literature.

The SPR test developed from the above ratio is

$$\frac{\beta}{1-\alpha} \leq B < (T_0/T_1)^r\, e^{-\left(\frac{1}{T_1} - \frac{1}{T_0}\right) V(t)} = \frac{P_1}{P_0} < A \leq \frac{1-\beta}{\alpha} \qquad (24)$$

If violated on left	If not violated	If violated on right
Accept H_0	Continue test	Accept H_1

The decision criteria are as noted.

A test could be conducted by taking one observation (one failure) at a time and computing the SPR at each observation to determine if a decision could be made at that time.

<u>Simplifying the Form</u>. Eq 24 is commonly manipulated to simplify decision making during the test as follows.

$$\ln B < r \ln (T_0/T_1) - \left(\frac{1}{T_1} - \frac{1}{T_0} \right) V(t) < \ln A \tag{25}$$

$$\frac{\ln B - r \ln (T_0/T_1)}{\frac{1}{T_1} - \frac{1}{T_0}} < - V(t) < \frac{\ln A - r \ln (T_0/T_1)}{\frac{1}{T_1} - \frac{1}{T_0}} \tag{26}$$

Since $T_0 > T_1$, $\left(1/T_1 - 1/T_0\right) > 0$. Defining $\left(1/T_1 - 1/T_0\right)$ as w for brevity and T_0/T_1 as k (which is commonly called the discrimination ratio, Ref 23:25)

$$\frac{\ln B - r \ln k}{w} < - V(t) < \frac{\ln A - r \ln k}{w} \tag{27}$$

Dividing by minus one and reversing the inequalities

$$\underset{\text{(Reject } H_0 \text{ side)}}{\frac{- \ln A}{w} + \frac{r \ln k}{w}} < V(t) < \underset{\text{(Accept } H_0 \text{ side)}}{\frac{- \ln B}{w} + \frac{r \ln k}{w}} \tag{28}$$

This relationship is commonly expressed as

$$h_1 + r s < V(t) < h_0 + r s \tag{29}$$

where

$$h_1 = \frac{-\ln A}{\frac{1}{T_1} - \frac{1}{T_o}} \quad , \qquad h_0 = \frac{-\ln B}{\frac{1}{T_1} - \frac{1}{T_o}} \tag{30}$$

and

$$s = \frac{\ln (T_0/T_1)}{\frac{1}{T_1} - \frac{1}{T_0}} \tag{31}$$

$$A = \frac{1 - \beta}{\alpha} \quad (2), \qquad B = \frac{\beta}{1 - \alpha} \quad (3)$$

Eq 5 given earlier is the commonly used normalised form of Eq 29. The work above is similar to the presentation by Epstein and Sobel (Ref 5:4) and later papers by Epstein. Handbook H108 (Ref 18) gives the above h and s values normalised to T_0 for many sequential test plans.

To check that the reject and accept sides noted above are correct, first it is noted that T_0/T_1 or k is greater than 1. Therefore, r ln k is positive.Therefore, the second term, r s, of each side of the Eq 29 inequality is positive. If α and β are both less than 0.5, which they must be, then A is greater than one and - ln A is negative. B is less than one, so - ln B is positive. Therefore, the right side, accept H_0, is larger than the reject side. Since H_0 is for a high T, it is expected that for the same number of failures, the V(t) must be high to lead to acceptance of H_0. Therefore, the decision criteria check.

Appendix B

Basis of A and B Decision Values

Wald (Ref 15:41) and the Columbia Statistical Research Group (Ref 14:B.05-.07) give in effect the following reasoning behind the A and B decision values and their relationship to the sampling risks. The liklihood ratio or SPR is set up as

$$B < \frac{P(\text{get sample got}/H_1 \text{ true})}{P(\text{get sample got}/H_0 \text{ true})} = \text{SPR} = \frac{P_1}{P_0} < A \qquad (32)$$

The A and B constants are independent of n, the sample size. A and B are combinations of the α and β sampling risks such that the probability of accepting H_1 given H_0 true is α and the probability of accepting H_0 given H_1 true is β.

If H_0 is true the denominator of the SPR should be larger than the numerator and the ratio less than one. Some B less than one can be selected such that if the SPR calculated for a particular sample is less than B, the H_0 hypothesis is accepted. If H_1 is true the SPR should be greater than one, and at some A larger than one a decision to accept H_1 can be made.

Say for a particular test the true SPR is greater than A. That is, the probability of getting any particular r events of n observations giving SPR greater than A, since H_1 is true, is at least A times as great as the probability of getting the same sample when H_0 is true. That is, the probability of getting a sample that makes the calculated

SPR greater than A is at least A times the probability of getting the same sample when the true SPR is less than B (H_0 true). Or, stated algebraically

$$\frac{P_1}{P_0} \geq A, \qquad P_1 \geq AP_0 \tag{33}$$

or

$$P(\text{get SPR sample} > A/H_1 \text{ true}) \geq A(P(\text{get SPR sample} > A/H_0 \text{ true})) \tag{34}$$

$$P(\text{accept } H_1/H_1 \text{ true}) \geq A(P(\text{accept } H_1/H_0 \text{ true})) \tag{35}$$

$$P(\text{reject } H_0/H_1 \text{ true}) \geq A(P(\text{reject } H_0/H_0 \text{ true}) \tag{36}$$

or, since

$$P(\text{reject } H_0/H_1 \text{ true}) = 1- \beta \tag{37}$$

and

$$P(\text{reject } H_0/H_0 \text{ true}) = \alpha \tag{38}$$

then, from Eq 36 $\quad 1- \beta \geq A \qquad (39)$

which gives the important

$$A \leq (1- \beta)/\alpha \tag{40}$$

Say that for another test the SPR is truly less than B. That is, any SPR calculated from the sequence of observations making up the sample should be less than B. Then,

$$\frac{P_1}{P_0} \leq B \quad , \qquad P_1 \leq BP_0 \tag{41}$$

where B is less than one so P_0 is more probable.

Then, the probability of getting a sample making the SPR less than B when the true SPR is greater than A (H_1 true) is at worst or is no larger than B times the probability of getting a sample giving

SPR less than B when the true population SPR is less than B (H_0 true),

or P(get SPR sample $<$ B/H_1 true) $\leq$ B(P(get sample SPR $<$ B/H_0 true)) (42)

or, P(accept H_0/H_1 true) $\leq$ B(P(accept H_0/H_0 true)) (43)

and substituting risks again

$$\beta \leq B(1-\alpha) \tag{44}$$

$$B \geq \beta/(1-\alpha) \tag{45}$$

The above A and B relations are those commonly used with the equality sign in SPR tests of all distributions. α is the producer's risk and β is the consumer's risk. Wald (Ref 15:42) and Epstein (Ref 8:2.77) indicate that for practical applications the A and B relationships using the equalities yield adequate approximations for the true sampling risks of the test derived from the relationships.

The effect of using an A that may be slightly too large (using the equality) is to increase the decision value for accepting H_1 and to spread the no decision range (see Eq 24). A decision to accept H_0 can usually be made as the V(t) value for acceptance of H_0 at some particular r is crossed by the event plot (see Figure 6). That is, there is usually no excess over the boundary or over the accept line for accepting H_0. Usually there will be an excess over the reject boundary since the event plot moves in discrete jumps of one r at a time along the abscissa in the usual SPR test plot. An event will seldom occur exactly at the reject line. The B value is usually exact then, but if the equality for B should be wrong the effect is to reduce the size of the B used from the proper size, and again the no-decision area is increased. Using the equality signs then tends to be

safe - that is the true risks are lower than those indicated in the calculations.

Appendix C

Obtaining an O.C. Curve

The O.C. curve gives the probability of accepting H_0 (that T_0 is true) given any true T. The abscissa in the plot is frequently given, as in this paper as, T/T_0. Wald gives procedures for obtaining the O.C. curves for SPR tests for other distributions such as the normal and and the binomial (Ref 15). Only the exponential is considered in this section.

Need for O.C. Calculations. It is assumed that few calculations for O.C. points as described below will be needed - Handbook H108 should supply curves for most purposes (Ref 18:2.5-2.24). If calculations are needed the procedures described in this appendix give good approximations. There is a good deal of literature concerning exact O.C. curves for the exponential. Such literature usually describes a good deal of effort. If production is relatively continuous or repetitive and if the quantities are large, exact probabilities and risks may be useful. For estimating the parameters of a test involving low quantities, particularly when the T expected and the T actually needed is subject to approximations that approach guesses, the value of exact O.C. curves is questionable.

Epstein and Sobel indicate that for exponential SPR tests, good approximations for the O.C. curve can be obtained from the following pair of parametric equations which have been slightly modified for this presentation. For convenience, they have been normalized to T_0.

$$L(T/T_0) = \frac{A^s - 1}{A^s - B^s} \tag{46}$$

$$T/T_0 = \frac{(T/T_0)^s - 1}{s\ (T/T_0 - 1)} \tag{47}$$

where s can assume any real value (Ref 5:4). For a particular s, some T/T_0 will be obtained from Eq 47. That T/T_0 value is paired with the value obtained from Eq 46 using the same s to yield $L(T/T_0)$. As many points of an O.C. curve as desired can be generated. An example follows.

For T_0/T_1 equal to 2 (note T_1/T_0 is 0.5), and α and β risks of 0.1, A is $(1 - \beta)/\alpha$ or 9 and B is $\beta/(1 - \alpha)$ or 1/9. From Eqs 46 and 47 the following combinations of s, T/T_0, $L(T/T_0)$ are obtained: 2, 1.5, 0.99; 1, 1, 0.9; 1/2, 0.828, 0.75; 0, ?, ?; -1, 0.5, 0.1.

<u>Sample Calculation for O.C.</u> It was noted that the points obtained from the above relations agreed quite well with equivalent points from the O.C. curves of Handbook H108 (Ref 18:2.15). The following is a sample calculation.

Using

$$s = 0.5$$

$$T/T_0 = \frac{2^{0.5} - 1}{0.5\ (2 - 1)} = 2\sqrt{2} - 1 \simeq 0.828$$

$$L(T/T_0) = L(0.828) = \frac{9^{0.5} - 1}{9^{0.5} - (1/9)^{0.5}} = \frac{2}{3 - 1/3} = 0.75$$

O.C. Approximation Points. If O.C. approximation points are necessary for the purpose of selecting an SPR test, five points should be adequate for most problems. Epstein suggests the following approximations be used since they are simple and adequate for most test selection use. A few selected sample calculation points have indicated that Handbook H108 points are in agreement with the points below, which leads to the inference that the handbook was based on these approximation points for O.C. curves.

Table II
O.C. Approximation Points

T	0	T_1	s	T_0	inf.
L(T)	0	β	$\dfrac{\ln A}{\ln A - \ln B}$	$1 - \alpha$	1

Appendix D

Obtaining Average Failure Number AFN and Waiting Time

Handbook H108 gives AFN values for many SPR tests (Ref 18: 2.63-2.65). If they are not suitable, the following relationships given by Epstein and Sobel may be used after the L(T) values of Appendix C are obtained (Ref 5:5).

$$\mathrm{AFN} \approx \begin{cases} \dfrac{L(T)\ \ln B + (1 - L(T))\ \ln A}{\ln\ (T_0/T_1) - T\left(\dfrac{1}{T_1} - \dfrac{1}{T_0}\right)} = \dfrac{-h_1 - L(T)(h_0 - h_1)}{s - T} & ;\ T \neq s \\[2ex] \dfrac{-(\ln A)(\ln B)}{\left[\ln\ (T_0/T_1)\right]^2} = \dfrac{-h_0\ h_1}{s^2} & ;\ T = s \end{cases} \qquad (48)$$

Usually an AFN curve approximated by values at a few critical points should be adequate. Epstein suggests the three points below. Terming T_0/T_1 as k

$$\mathrm{AFN}(T_1) = \frac{\beta \ln B + (1 - \beta)\ \ln A}{\ln k - (k - 1)/k} \qquad (49)$$

$$\mathrm{AFN}(s) = \frac{-(\ln A)(\ln B)}{(\ln k)^2} \qquad (50)$$

$$\mathrm{AFN}(T_0) = \frac{(1 - \alpha)\ \ln B + \alpha \ln A}{\ln k - (k - 1)} \qquad (51)$$

(Ref 6:11)

To obtain the expected waiting time the equation below from Epstein and Sobel may be used (Ref 5:6,8). If the test is without replacement Handbook H108 gives a different relationship (Ref 18:2.58-2.59).

$$E(t) = \frac{AFN \cdot T}{n} = \frac{(\text{Failures})(\text{Time/fail})}{(\text{equipments})} \qquad (52)$$

Where E(t) is the calendar time, AFN is the average failure number, T is the true mean time between failure, and n is the number of equipments on test. Manipulating this somewhat

$$\frac{nE(t)}{T_0} = AFN \frac{T}{T_0} = \frac{V(t)}{T_0} \qquad (53)$$

where V(t) is the sum of the test time on all items. Therefore, for a particular test with replacement the AFN values can be obtained from the above calculations or from Handbook H108 if appropriate, and the expected waiting time V(t) for various possible T true values calculated easily.

For example, to obtain the $V(t)/T_0$ values for various T when using Test C-11 of Handbook H108, Table 2A-1 gives a ratio of 0.512 for T_1/T_0 (Ref 18:2.2). It should be noted that this is quite close to the AGREE Task Group 2 test (Ref 13:88). Table 2D-1(c) of Handbook H108 gives at T equals T_1 (θ_1 in the table) an AFN (E_r in the table) of 9.7. That is, if T equals T_1, 9.7 failures per test can be expected on the average.

Since T/T_0 is T_1/T_0 and

$$V(T_1)/T_0 = AFN(T_1/T_0) = 9.7(0.512) \cong 4.97$$

If T is T_0, T/T_0 is one and

$$V(T_0)/T_0 = 6.2(1) = 6.2$$

Appendix E

Estimating $V(T_0)/T_0$ Versus T_1/T_0 Relation

The sketch of Figure 16 indicates that an exponential relationship is a good possibility. If true, the equation will be of the form

$$y = e^{ax + b} = Ce^{ax} \tag{54}$$

$$V(T_0)/T_0 = e^{a(T_1/T_0) + b} \tag{55}$$

This is the expected test time in multiples of T_0 for a test with ratio T_1/T_0 if T_0 is truly the lot T. Then,

$$\ln y = ax + b \tag{56}$$

where

$$\text{Slope} = a = \frac{\ln y_1 - \ln y_2}{x_1 - x_2} \tag{57}$$

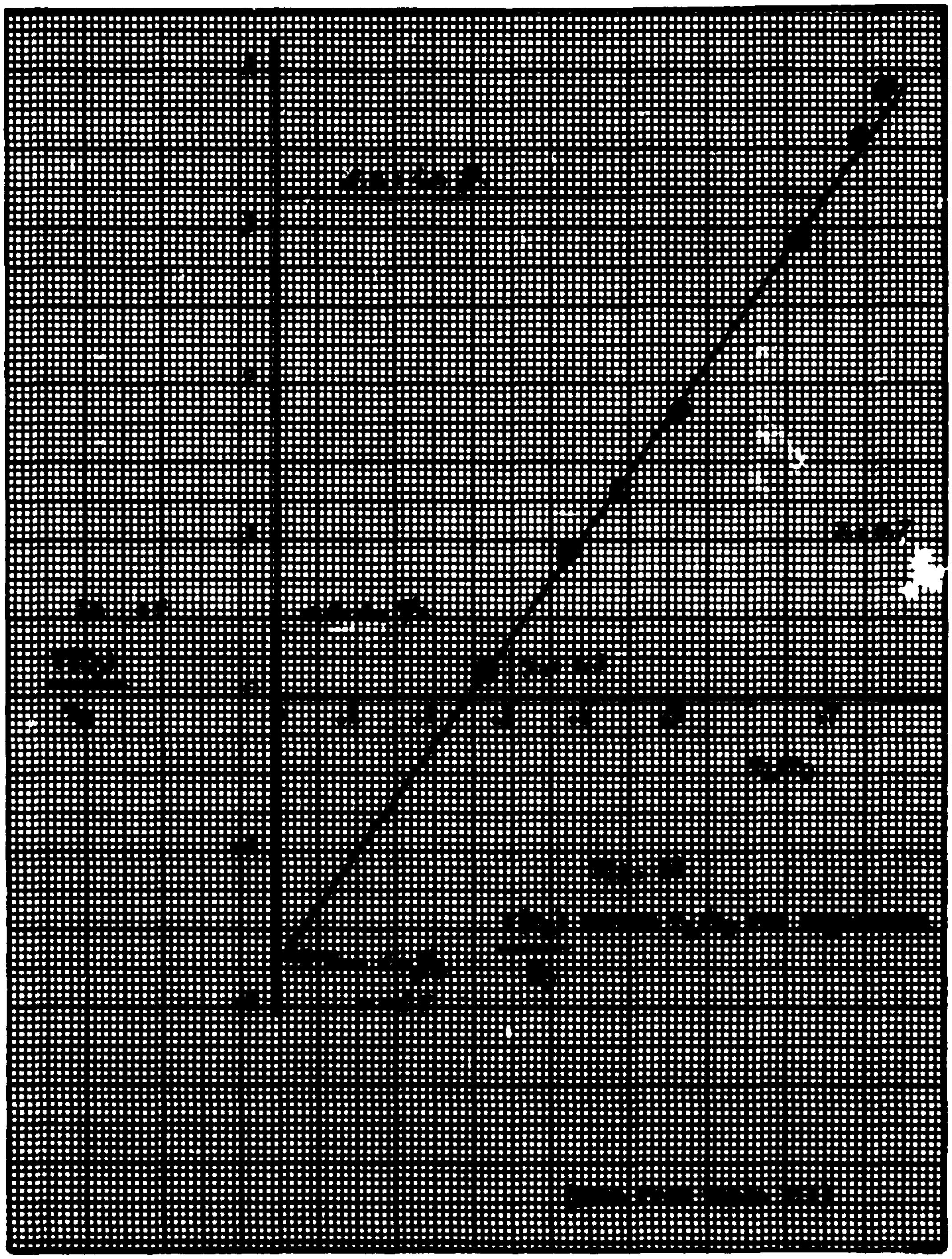
$$b = \ln y_b \tag{58}$$

and

$$V(T_0)/T_0 = (V(t)/\text{given } T = T_0) \cdot 1/T_0 \tag{59}$$

The data of Table III were obtained from Handbook H108 initially and calculations made as described in Appendix D.

When $V(T_0)/T_0$ data from the table is plotted on the log axis of semilog paper against the T_1/T_0 ratio for the related test, if the relationship is exponential, the plot will be a straight line. Figure 18 shows that such a relationship holds quite well. A formal



regression analysis could have been run, but a visually placed line is often sufficiently accurate, particularly when the points deviate but little from a line. The error calculations indicate that within the T_1/T_0 range of approximately 0.26 to 0.72 the error is less than plus or minus 7 %.

It seems appropriate to note here that the Handbook H108 code C-11 and C-15 tests noted in Table III are similar test plans to those of AGREE Task Groups 2 and 3 and of MIL-R-26667 Methods 1 and 2 (Refs 13, 19).

Table III

Tabulation of AFN, $V(t)/T_0$ Versus T_1/T_0

Code (H108)	T_1/T_0	AFN(0) yields $V(0)/T_0$	AFN(T_1)	$V(T_1)/T_0$	AFN(s)	$V(s)/T_0$	$V(T_0)/T_0$ AFN (T_0)	ln $V(T_0)/T_0$	$V(T_0)/T_0$ calc	Error
C-4	0.261	0	2.9	0.756	2.7	1.28	1.2	0.1822	1.13	-0.0534
C-7	0.370	0	4.8	1.78	4.9	7.86	2.5	0.9155	2.44	-0.024
C-10	0.438	0	6.7	2.94	7.1	4.57	3.8	1.335	3.92	0.0316
C-11	0.512	0	9.7	4.97	10.8	7.59	6.2	1.823	6.6	0.0645
C-15	0.666	0	24.3	16.2	29.2	23.6	18.5	2.92	19.45	0.0512
C-17	0.743	0	43.9	32.6	54.7	47.0	36.0	3.58	23.95	0.161
C-18	0.774	0	48.2	37.3	73.6	57.0	49.1	3.895	30.2	0.163

Note: AFN(T_1) signifies the AFN when T_1 is true T.

$V(T_0)/T_0$ signifies the expected sum of test time in multiples of T_0 when T_0 is the true T.

<u>Sample Calculation</u>.

Slope,

$$a = \frac{3.2 - 0.4}{0.7 - 0.3} = 7$$

Intercept,

$$b = \ln y_b \simeq -1.7$$

Regression equation

$$\ln y = 7(T_1/T_0) - 1.7 = \ln e^{\left[7(T_1/T_0) - 1.7\right]}$$

$$\therefore\ y = V(T_0)/T_0 = e^{\left[7(T_1/T_0) - 1.7\right]} \qquad (11)$$

or

$$y = e^{7(T_1/T_0)}\ e^{-1.7} = \frac{e^{7(T_1/T_0)}}{e^{1.7}} = \frac{e^{7(T_1/T_0)}}{5.474}$$

$$\therefore\quad V(T_0)/T_0 = 0.183\ e^{7(T_1/T_0)} \qquad (11a)$$

Error,

$$Er = \frac{V(T_0)/T_{0\ calc} - V(T_0)/T_0}{V(T_0)/T_0}$$

At T_1/T_0 equal 0.666

$$V(T_0)/T_{o\ calc} = e^{7(0.666) - 1.7} = 19.45$$

$$V(T_0)/T_0 = 18.50 \quad \text{(From Table III)}$$

$$Er = \frac{19.45 - 18.50}{18.50} = 0.0512 \ \text{ or } 5.12\ \%$$

Appendix F

Risk Considerations

Table IV
Considerations in Choosing β, Consumer's Risk or T_1

Level	β low or T_1 high	β high or T_1 low
Consumer wants Level -	If: 1. Application critical 2. Items are for GFAE to a system contractor 3. Producer's reputation questionable 4. Past performance poor 5. Facilities inspection causes doubts 6. Producer's procedures questionable in: design, quality control reliability plans, organization, knowledge program data feedback, use other 7. Producer's attitude poor 8. Producer does not seem to understand problem	Since: 1. Tends to reduce test time If: 2. Must obtain even low reliability items soon 3. Estimated costs of T_1 items not excessive Consider: repairs expected in use lost time when fail loss of a/c or missile or other loss of mission of supported system redundant items needed to assure function when needed added personnel, spares, facilities due to T_1 4. Producer's reputation good
Producer wants Level -		Since: 1. Tends to reduce test time 2. Fewer marginal lots rejected

Table V
Considerations in Choosing α, Producer's Risk

Level	α low	or α high
Consumer wants Level -	Since: 1. Will accept more lots near T_0 If: 2. Items needed soon 3. Impact of delayed delivery of good items will be high 4. Producer must add too big a margin to his price to cover possible losses due to rejection of good items	Since: 1. Tends to reduce test time 2. Forces producer to set target T higher than T_0 to reduce lots rejected If: 3. Information indicates that the true T will probably be about T_0, but the variation from unit to unit will be large
Producer wants Level -	Since: 1. Will reject fewer lots near but below T_0 If: 2. Impact of rejection will be high 3. Higher risk will force bid price too high to be accepted	Since: 1. Tends to reduce test time

Appendix G

When to Use SPR Tests

The Columbia Statistical Research Group document gives rules for deciding when an SPR test should be used rather than a conventional test (Ref 14:1.08). These rules are included in essence in this appendix. If an SPR test is used to test a characteristic of items where the unit of test is one item, a reserve quantity of those items is required in excess of the ASN, the number that will be needed on the average for test. There is a probability that a few SPR tests may give a plot wandering in the continue-test zone of the decision chart. In tests of mean time to failure, the unit of test is not an item but a failure. Therefore, if the items on test can be repaired and returned to test after they have failed, those items represent a large number of potential units of test on reserve far greater than the AFN, the average number of failures expected. Therefore, the cost of reserve units for reliability tests is generally low. This influences Table VI below.

<u>Advantages of SPR Tests</u>. The following advantages of SPR tests and the points of Table VI should be considered in selecting a test.

a. Duncan is one of several references indicating that for the same sampling risks (in effect the same O.C. curve) an SPR test will require less inspection and waiting time than a conventional test. (Ref 4:156).

Table VI
Test and Unit Cost Influence on Test Choice

		Cost of getting units for test and for reserve: High	Cost of getting units for test and for reserve: Low
Total Test Cost	High	Sort of test depends on relative cost For MTBF tests, SPR probably best	Use SPR Minimizes number of units tested
Total Test Cost	Low	Preassigned time or number of failure test probably best Cost of reserves should offset SPR test saving in units tested	Sort of test depends on relative cost If time required for preassigned time or r test is not objectionable, such probably best If waiting time is valuable, SPR test probably best

b. Inspection charts or tables are easily drawn once T_1, T_0, α, and β are selected. The charts or tables can be easily changed during a production run if data indicate that the test should be changed.

c. The reduced waiting or test time when true T is greater than T_0 is a built-in incentive for the producer to deliver good equipment. This, though, may not always be true. If the test costs for a particular item are low and if the trouble and costs to the contractor due to rejected lots is low, the incentive may be negligible.

Disadvantages. The following should also be considered.

a. Measurements must be made quite frequently and much more data must be recorded than in the usual conventional test. In some conventional tests the number of measurements can be greatly reduced.

b. Reserves of units of test must be available. As mentioned above, this point is not particularly appropriate for MTBF tests.

Appendix H

Test Considerations to Verify for Checkout Equipment

Before the physical and the statistical provisions of a test plan for reliability are released or approved it is suggested that at least the conditions be verified as true. The lists can hardly claim to be complete, but they are intended as a guide to lead to more searching questions for tests of checkout and other equipment. It is no easy task to clearly state test objectives - it is even more difficult to attain them.

Physical Test Conditions to Verify

1. The expected service environment is simulated or provided for.

2. Failures have been clearly defined - both catastrophic and drift.

3. Two different decision test plans have been considered. One for normally undetected drift or tolerance failures which may give a high probability of causing complete loss of a combat vehicle. Another for catastrophic but normally detected failures which may delay mission start.

4. The impedance the item sees or the functions it responds to or performs adequately simulate expected use.

5. The accuracy of measuring equipment used is adequate.

6. The item is cycled through its various modes of operation in

a manner proportional to its expected use.

7. Measurements of performance are made frequently enough.

8. The stimuli to which the item is expected to be most susceptible (for example, overvoltage, undervoltage, low temperature, prolonged operation, electromagnetic interference) are applied and tests of response made.

9. Characteristics that have been troublesome in similar items are specifically covered in the test plan.

10. If environmental tests of the complete item are at reduced levels due to size or cost, at least the assemblies identified as critical or common to many modes of operation are being tested at realistic environments.

Statistical Test Conditions to Verify

1. β and T_1 are adequate.

2. T_1 is not so low that redundancy or standby units will be needed at the use site to insure that the function is available when needed. If T_1 must be low due to technical or other limitations, the need for redundant units should be consciously considered.

3. T_0 agreed to by the contractor is not so much greater than T_1 that the test is apparently set at much too low a T_1 level.

4. T_0 is not greater than the true T attainable. If the true T is between T_1 and T_0 the test waiting time may be high.

5. The number of items available for test is sufficient to give a reasonable assurance that a test decision will be made within calendar

limitations on delivery.

6. There is reasonable assurance or expectation that the test items will be representative of all equipments accepted or rejected by the test.

7. It is reasonable to assume that the exponential distribution will hold or at least there are no negative indications.

8. Malfunctions chargeable as test failures are clearly defined.

9. Failure data that can be censored are clearly defined.

10. Estimated test costs are a reasonable fraction of the total contract funds.

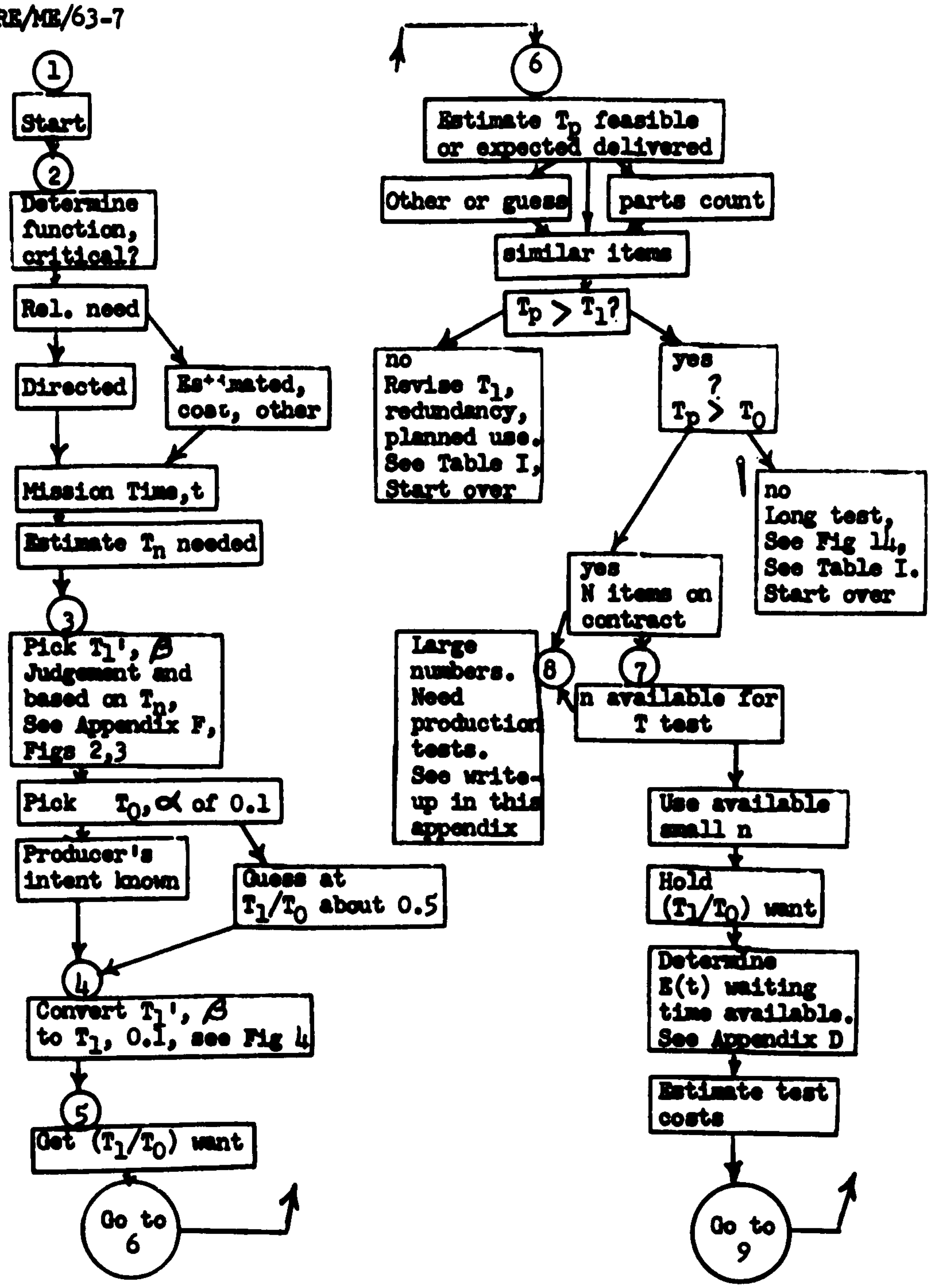
Appendix J

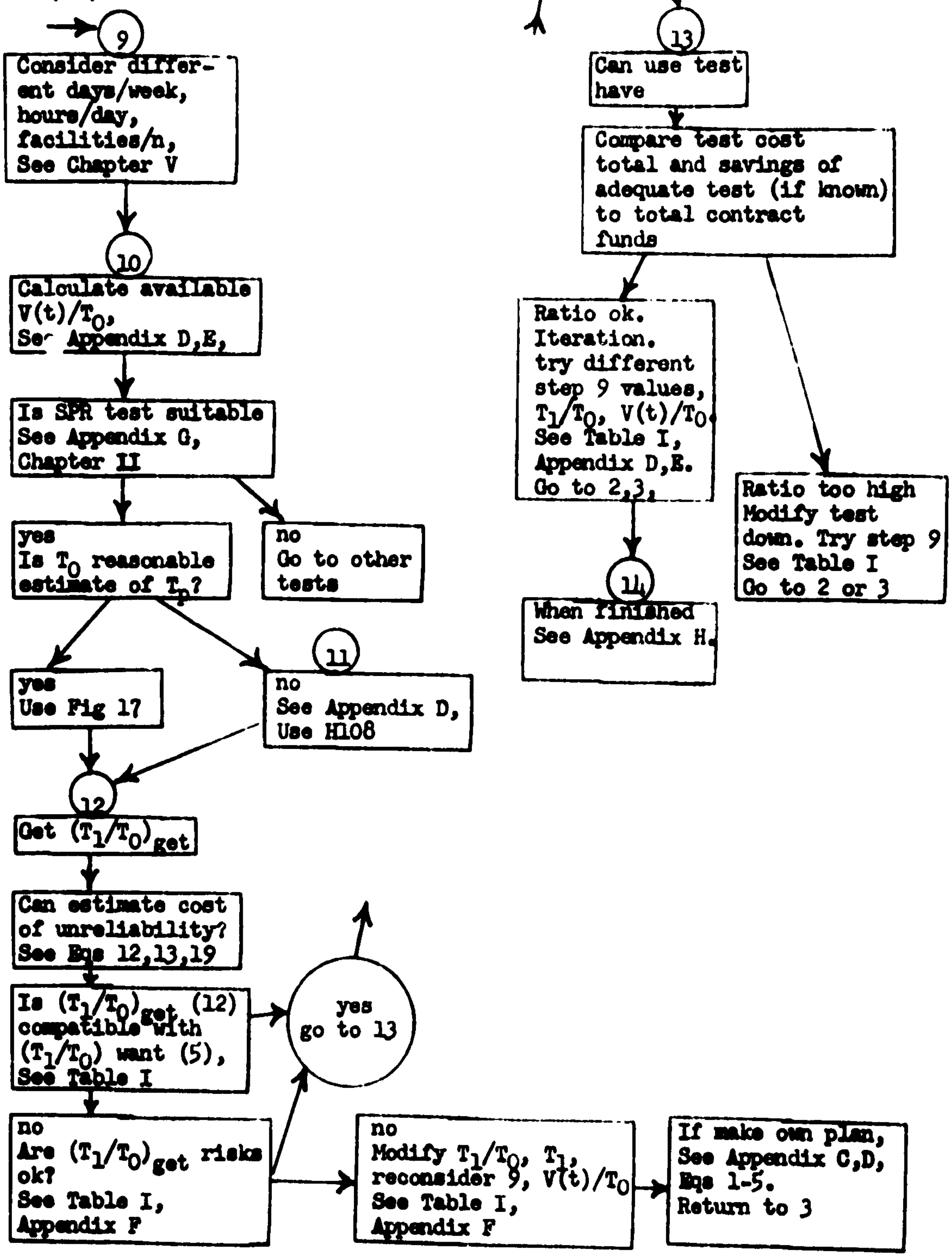
Flow Chart for Possible Test Selecting Procedure

The flow chart on the following pages cannot include all the details that make reliability work frustrating. The procedure is not elegant. A mathematical method of attack as mentioned in Chapter V would certainly be preferable. However, lacking such a time saving device, something on the order of the chart in this appendix is at least a starting point. Its success would depend on the persistence or stubborness of the person working on the problem.

The approach here is primarily that of a reliability engineer being requested to select a reasonable test plan for a contract for checkout equipment or any other equipment that could reasonably be assumed to follow an exponential distribution of failures.

In step 8 in the chart when the quantity on contract and the number of specimens or items available are large, production sampling plans such as AGREE (Ref 13:167) or other should be applied. Those methods are not covered in this paper.


1
Start
2
Determine
function,
critical?
Rel. need
Directed
Estimated,
cost, other
Mission Time,t
Estimate T_n needed
3
Pick T_1', β
Judgement and
based on T_n,
See Appendix F,
Figs 2,3
Pick T_0, α of 0.1
Producer's
intent known
Guess at
T_1/T_0 about 0.5
4
Convert T_1', β
to T_1, 0.1, see Fig 4
5
Get (T_1/T_0) want
Go to
6
6
Estimate T_p feasible
or expected delivered
Other or guess
parts count
similar items
$T_p > T_1$?
no
Revise T_1,
redundancy,
planned use.
See Table I,
Start over
yes
?
$T_p > T_0$
no
Long test,
See Fig 14,
See Table I.
Start over
yes
N items on
contract
Large
numbers.
Need
production
tests.
See write-
up in this
appendix
8
7
n available for
T test
Use available
small n
Hold
(T_1/T_0) want
Determine
E(t) waiting
time available.
See Appendix D
Estimate test
costs
Go to
9


9
Consider different days/week, hours/day, facilities/n, See Chapter V
10
Calculate available $V(t)/T_0$, See Appendix D,E,
Is SPR test suitable See Appendix G, Chapter II
yes Is T_0 reasonable estimate of T_p?
no Go to other tests
yes Use Fig 17
11
no See Appendix D, Use H108
12
Get $(T_1/T_0)_{get}$
Can estimate cost of unreliability? See Eqs 12,13,19
Is $(T_1/T_0)_{get}$ (12) compatible with (T_1/T_0) want (5), See Table I
yes go to 13
no Are $(T_1/T_0)_{get}$ risks ok? See Table I, Appendix F
no Modify T_1/T_0, T_1, reconsider 9, $V(t)/T_0$ See Table I, Appendix F
If make own plan, See Appendix C,D, Eqs 1-5. Return to 3
13
Can use test have
Compare test cost total and savings of adequate test (if known) to total contract funds
Ratio ok. Iteration. try different step 9 values, T_1/T_0, $V(t)/T_0$ See Table I, Appendix D,E. Go to 2,3,
Ratio too high Modify test down. Try step 9 See Table I Go to 2 or 3
14
When finished See Appendix H.

Vita

Ben Norman Theall was born 22 November 1925 in White Plains, New York, the son of Benjamin Harrison Theall and Florence Gertrude Theall. After graduating from White Plains Senior High School in 1943, he served in the U. S. Army, primarily Infantry, as an enlisted man from 1944 to 1946. After his discharge, he enrolled in New York University and in June 1951 graduated with the degree of Bachelor of Electrical Engineering. In 1951 he was commissioned a Lieutenant in the Signal Corps, USAR, and was honorably discharged in 1954. After his 1951 graduation he was employed by the General Electric Company in various engineering assignments. He entered the Civil Service in 1956 at Aeronautical Systems Division (then Wright Air Development Center) working in specifications and standards for electrical and electronic parts and later in standardization of various electrical and electronic parts and equipment. In 1960 he took the state examinations and was registered as a Professional Engineer in the state of Ohio. Before entering the Air Force Institute of Technology he worked in the area of reliability and maintainability of equipment assigned to the Directorate of Aerospace Ground Equipment Engineering.

Permanent address: 4339 Kitridge Rd.
Dayton 24, Ohio

This thesis was typed by Mrs. Jacqueline P. Theall.